FIGURE 1A

The Small Island: 10848bp (SEQ ID NO: 1)

TTTTCGAAGGGGCGAGTTGCACAGGCCTAGGAAGGGGTGTTGCACCTTTTACTTGTCCAGCAGCCGTCTGGGCCGAACAT
TAGGCCCGCATTGACGTGATGTCAATATGGAATAGCCCGCCTATTTAGGTGGGCTTTTCTGTTTTTCAGGCCCTGAAACC
ATTCTGCAGGCGTAGCAATGTGCCCTCGACGTAATCAGCGATTACAGAGGGTACTGCGGTTTCGGCCTTGAAATTGCGGG
GCTCTGCCAGCATGCCGGCGCTGGAATCCAGGCGCTCTCGGTAATTCCAGATTTTTCCTACAGGCCTAGCGAACTGCTGG
GAGGCGCAGCTACCTTTCGTCCGGTAGCCGGCCCCTCAGTATCGGTATCGCTAGGCGAGACTGCTTCGGGCGGTAGATCC
CCATCCTCCAGATCCGCTTGTTGCGGTGGAAGCCGCGGAGGCCGTGGCGGCACGGTTCGATAGGTGTTGAGCCATCAGAG
TGAACAGCACGCTTTGCTCGAAGGTGTCAGGTGGCTGTGGTAGCGCTTCTTCCCACGCGGTTGCATCTCCCTGAATGCCT
GTCCCTGGTGTGTGCTCCCTTGGTTCTGCAGGTTCTGTCCTAGCCTCCTGTGGCGGCGCTTCACCCATTCGAGCGAGAGG
TTGCCCGCATGGCTTCGGTTTGGCAGAGGCTTGCGGGGAGCCGCTGTAATCCACCTTCCTGACGTTCTCCAGCGCTCCTT
GGGCCAGGCTTCTGGCATGCAGCACGGCACGCATGTGGTCGTAGTCGGCCTCGATGTAGCTCATGGTGGTCTGGATATTC
GAGTGGTTGAGCAGGCACTTCGTGAGGTGGATGTTCCGCTCGGGTGCCTTCATCAAGTCGGTGGCCAGGGTGTGCCGGAA
ACGGTGCGGAGTCATCCGCACCCCAACCTTCTCGGTCAACTTCCGGTACATGGCTTCGACCTGGTCGGAGTTCATCACCT
TGCTCTTGTAGTGCGGTGAGAACCGGTTGACGTTGAACAACTGGTCGTCATCGGCGAATCCGGCTCTATCGGCCTCCTGC
AGGAGCCGCGATAGGTGAGGCACAAGCCCCTCCGTTATTGGCACTACGAACTCTTTATGAGTTTTCTCTGTCTCGCCGCG
GATGAGGATCAGTTGATTTTCCCAGTCGATGTCGCGCTTGCGGATGCACAACAGCGCATTCAACCGGATGCCGGTGAAGT
AGAAGACCTCAAACGTGCAAAGCCAGAACCAGGCGGGAGTGATGCGTGCGCGATCGCCAGTGCAGCGCTCGGCGCCGACC
TGCATGTTGAGCCAATTGCGGGCGCGCAGGATGGCTTCGGCTGCGACGGTTTTGCTTGCTCGCCTGGGGGGGATGACGGT
GGTCTTTCTGAACGGGTTGACTTGGGAGTGTGTCACCAGCTCGTGCTCGATGGCATAGCCCCAGATCGTTCGCAGATGAT
TCGAATACGTGTTCCAGCTCCGCTTCGACAGGCCTTGTTCCAGGACCTTGCGCCGCCATCCCAGCACAGACCTGTGGTCC
ACCTCCTGTACGGTTGCCGTAGGGCCGAAGTGCTTGAGCAGCGCCTTGGTCGCGGCGCGGTAGATCTTCGCGCTGGCTTC
TCGGAGATCGTGCGCGAAGATGTACTCCTCGGTGAGCTGCTGCGGCGTCATTCCACACCTCCTTCGGCATCGGTGATGAC
CGTGAGGCTTGGGTTGTCCAGAGGCTGCACAGGGAACAGCAATTTGGGATCCTGGAGCAGGTAGGCCTTGAGCTCTTTCG
TCTTGCGAGGACCAGAAACCTTGATGGTCCAGATATTCAGGTTCTTACTGGTCTTTCGGTGAAGACCCTGTTTCTCGAAC
GCGCGCTGCACCAGCTTCCAGCCGGCGGTCTCCTTGGCCTGGGCCAGCTTTTCAACCTCCGGATGCTCCTGGACATAACG
CTTGAAGATTCCTGGCGTGACCAGCATGGCGGTCCCGTCGACGGTATGCACCAAAGCCTTGGTGTCGTTGATGAACAGGC
GACGGGTCGCGATACCAGATTTCATCCAGCCAATGAATCCCTGTCCCAGATCAGTGCTTCTGCTAGGCATAAATACATCT
TCAATTGCTTCAGGAGCGCAATCTGTAGTGGTCAACCCATCCCGGACAGTGGGTTGAGTTTTTCTTCGGCCACCGCAGGT
GGTAACCCGTCGTTGAATTGATGCGGCCTGATCCAGTTGTAGCGGTGCATCAAGTAATGACTGATGTCCCGTTGGGCCTC
CTGCGCCGTCAGGTAACCCGTTGACGGGACCCACTCCGACTTCAGACTGCGGAACAGGCGCTCCATCGGCGAGTTATCCC
AGCAATTCCCCCGACGGCTCATGCTCTGCTGCATCCGATAGCGCCAGAGCCGTTGCCGAAACAGGCGGCTGGCGTACTGG
CTGCCCTGGTCTGAATGGAACAGCACCTGCTGTGGCCTGCCGCGCTGTTCGTAGGCCATGTCCAGGGCCTTGATCACCAG
TTCGGCATCCGGCTTGGCCGAGAACGCCCAGCCGATCAGCAGATCCAGCACCGCGGCCAGGTAGTGCCAACGGCCTTGCG
CCCAGACGTAGGTGATGTCGCCACACCACACCTGTATGGGATGCTCGGTCGCGAATTCGCGGTTCAGCCGATTCGGGATA
TCCGGCCGCTCAACCGTGGCCTGTTTGTAGGCGTGCGAGCCCGGTTGCTTGCTGACCAGGCCCAGCTCACGCATCAACCG
ACGCACTCGGAAACGGCCGATGGTCACGCCCTCTTCGCGCAGCATGCCCAGAATGCTGCGGCTGCCGGCCGAGCCCCGAC
TCTGGCTGAACAACTGGTTGACTTGGCTGCGTAGCGCCACGCGGCGAGCATCGACACGCCGCCGTCGAAGACGGTGGACG
TAGTAGCAAGACCGCGCCACATCGAAAGCTGAGCAGACCACTTCCACCGACTCCTGCTCACTCAACTGGTCTATCAGCGC
GTACGATCGAGTTCGTCCGACATCAAGAGAGCGGTAGCCTTTTTTAATATCGCTTTCTCCCGCTCCAATCGGTTGATCCG
GGCTTCCAGCTCCTGGATCTTTTGCTGCTCAGGCGTCAACGCCTTGCTCTTCGGGGTCACACCCTGGCGCTCCGCCTCGA
GCTGCTTCACCCAACGGCGCAAGGCCGAATCCACCACCCCCAGCGAACGGCAGGCGTCGATATGGCTGTAGCCTTGGTCC
AACACCAGGGCCGCGGCCTCTCGTTTGAACTCGGCGGAAAACGTACGTCGTTGCTTGCTCATCAGACACCTCTTTCACGG
CGAGGATTCTCGCCTAAATCGGTGTCCGGGATCAGTAGACCACTACACTGCCGAAACAGTTATTTTTGTTGCTAAGAGGC
AGGAAGGCTGCGCTGAGATTCCTACTAGTGAATTAGATAGTTGTTGTCTTCTGAAACAGAACTGAAGCGAAATTGGGGGT
AGGGTTTTTCTAGGTGAAGGGTAACTCTTGAGAATTACAAAGGTGTGACGCAAATGTTGTATTTTTCTTGCAGTATGAAG
ATGGGTGGTTGGGTCGGATATAGGTACTTCTCTCTATTTTCTTTAATTGCTCTCATCTATGGGTGTGTCGGTGGTGGAGG
TGGATCGGATGAGATTGGGCAGCACTGCTTTGAGAGAGAGCAAAAGCTTTCCGGAGTTAATGATAATGAAGAGGGGAGTG
TGAGGTTGAATCGGCTGAACTGCGATCCAATTGAAGGTCGTGTTCTTGAATCAGAGAAGCTGATAAGAAAGCCGCCCAAT
GAGCTGGGTATTCACTGAGTTAATAATTTTTAGCCTGGAGAGGTCTTTTTGAAAATGAAAAAATCACTTGTTATGTCGGC
TGTGCTTTTGGTGGCTAGCAATTTCGCGTGTGCTGATGAGGGCTCAAATGATGGAAGTGAGATATGTCGGGCGCAGGGTG
GAGTTGAAATAACAAGTCTGGGGGAAGTCTCAAAGGGTGTGGATGTTGAAGATGTTGTAGTTGTTCGATTCTTCCAAGT
AATATGAAGTCGAGTCAAAGAGCGCCTACACTCCCTCCTCTGCAAAGGATGATCATTTCGGCAATGCCTTCACCAGGAAC
GGTCACTGTTTCTGCCAGCGGAGATAGGAAATTTACAACATCTTGCCGGGCAAATCTTTATGCTCCACGTTATGCCAATT
TCTATCCAGACGGTGTTAGCAGGGGAACATCAGATCTACGATGTGTTGGTTACAATACACCCGGGAATTCATCTCAAGGG
TGTAATGTGTCATGGGACGGCCCGACCGACATTCAATTGGGTGTTGAGCCATATGGCGGATCTGTTGTTGTAAACTACAG
TTGCACTGCATTCAAAACAACGATTCCAGTGATAATGAGCTACAGTTATCGTGATGGGCGGGCAGTGTATGGCGAGGTCC
AGAATGTGTCAGGAATAATAAATGTGGTTTTGAACTAAGGATGCGCCTTAGTAGGATTTTATTTCAAGGTGTTCAGCTTG
TGAATGTGCTGCAGGGCACTCTATAGAGTGCCCTGCATAGGCTGGAAGTCGTTTGGCTGAGGGCTAGTCTTTCGAGTTCG
ATATGGCACGTTGATGAAGGCTGGTAAAAAGTCCTGTCTGTGTGGGTGGCAGGCGAGTGGCGCGTTCAGACGCTAATGTG
GAGTTCAAGGTAGGGTTCCGACGCCGAGGATCCTCTTTGGGTGGAGTCATTTGCTGAGCCATGCAGTTATAAGACAGGCT
ACGTGCTCCATTGACTTGGACATCGATACACGAAGATTCAGATTGTTCCGAGTGTGACTCTGCGGGCGGGGAGGCAGCAG
CCTGCTGTGCCAAACCAACTATAGGCAACAAGAATATAATTCGAAGAATTTTAATAAGCATTCCAATCTCCTTAACCCTG
AGGGTTGGATGAGACTCAATAGTAACCTCAAGTTATGAGGAAAGATGAACCGCCTATCGTTTTGTAGTCTTGGTTAGGTC
AGGCACACGCTAGTGAGTAGTACTAAGAGTAAGCCGATAGCCAGGGGGCGTGGTGGCCCATTTGGGGAAGTGATGAAGAG
GTGCGGGCTTGTACCGGTTCGAGGAAGGAATAGACAGCAGACAGGATCGCTTGCGATGGGGCAGCAGGAAACCATCAGCC
CGTCCGTATCCAGAACTGCTGCTTGCAGCGTTAGGGGTGACTCCCTCATGCCCTAGAGGGCATAATCCCTGTATGGAGAT
GGCTAACTGATATTAAGGAGTGGTCGAAGTCATGGAACGCTTGCTCGAGAGCATTTACATCAATGCCCGGCCGGCGATGG
AGTTGAGGCTTAGCCTCACCAGCTCCGGCCGCAAGAGAATGGTAAAGATTGTGGATGGGGAGGAGGTCGAGGTTCTGCCA

FIGURE 1B

GGTGAAGTGCAGGGCATCCTGGAGGCCCAAAAGAGGGATGTTGGAATCCTCGCCGACTTCTTAGCCAAGAGTCTCGTGGC
GCGACGCTAGAGATCCTTGGCTAGCCAGTCTATCGCCGACAGCAACGTGGTCTCGTCGGCCATGCCTTCAAGCTTCACGC
TTCGTCCTGTCTGGTTGTCCCGGATGACGAGGGTCGGGGTCGCTGTAATGCCCTTCGACTGTGCGGTGTCGATGTTCGTA
AGCGTCCGGGGAAGGCACATTGCGATGCGCTAAGCTGTAATCCTTTAGGATTGAGGTCTCAGCATGGAATGCCACGTTCG
TCCCGCCACGAGCAGAGATGCAGCAGCGATAAGCTGCGTAGTTATAGCCGCCCTGCGTGAGTCAAATTCACAGGACTATC
CGCCTGATGTGATCGCTCAGGTTGAGCAGAGCTTTTCTCCTGAAGCCATCACCACACAGCTTACGAAGCGTAGGGTCTTC
GTAGCCTTATTGGGCGAAAACATTATTGGCACTGCCGGTCTCGACGGTGACGTCGTCAGAAGTGTTTTCGTTGACCCAGC
TCACCAGAAAGGCGGTATCGGGCGGCATTTGATGGATGTCATTCATACAACTGCTGCCAGCGCGGGAGTTGGAGCTGTAC
GTGTGCCATCGTCGATTACAGCTGAAAGGTTTTATACCGCATTGGGTTATCAGAAAATCCGCGACGAGTTTCATGGGGCG
GAGCGCACCATCGTTATGGAGAAGCGGCTGTAGGATTATTCCAGACTATTTACCCACTTATGCGGCAACAAATCAGTGAT
CTCACTGGCCCGCTGCGTGGGCAGCCGCGTCAGCACATCTTTGAGATAGGCATACGGATCATGCCCATTCATACGCGCCG
ACTAGATCAGGCTCATGATTGCAGCCGCCCGTTTACCACTGCGCAGCGACCCGGCAAACAGCCAGTTCGAACGCCCGAGC
GCCCATGGCCGTATTTGGTTCTCGACCTGATTGTTATCGATGGGCACAGCCCCATCGTCCAGGTAGCGCGTCAGCGCTAC
CCAGCGTTTCAGGCTGTAATCGAGGGCTTTGGCCGTGGCTGATCCATTGGGCACCAGGTCTCGCTGAGCCAACATCCAGT
CATGCAGTTTTTTGAGGATCGGCACCGCCAATTCTTGTCGTATTCGCCAGCGCTCTTCATCACTCATGTCCCGCGCCTGA
CGTTCGACCTCGTACAAGCCGCTGATCGAGTGCAGGGCCTGTTCAGCCAGCTGACTTTTGTTCGCCACGTGCAAATCAAA
GAACTTGCGCCGGGCGTGGGCCATGCAGCCGATTTCAGTGATGCCTTGTTCGAAACCGGCTTTGTAGCCAGCGAAGTCGT
CGCAGACCAGCTTGCCGTTCCAGTCACCCAGGAAGTTGCGCGCATGTTCGCCAGCACGGCTTGGGCTGAAGTCGTAAACC
ACCGCTTTGAGCCCTGAAAACGGCGTCGTGCTGTACGCCCAGACATAGGCCCGGTGGGTTTTCTTCTCGCCTGGCGCAAG
CATTTGCACCGGTGTTTCATCAGCGTGGATCACGCCCTGGTTCAGCACGGCTTCACGCAGTGCATCGACCAGTGGCTGAA
GCCGCACGCCGGTTTGTCCGACCCACTGCGCCAGGGTCGAGCGAGCAATTGCCAGCCCGGCGCGGCCAAAGATTTTCTCC
TGCCGGTACAGCGGCAAGTGATCGGCAAACTTGGCCACCATCACGTGGGCCAACAAACCTGCGGTCGGGATACCTTTGTC
GATTGACCTGCCCCCCGATTTCAGTACCACGCTGAACTTAGTAGAATCCGTTTTCCAAGCAGGAGACGGCAGTGCAGAAG
CGTTCTTCTCGACTTCCTCAAGCGGGTCGAGGCTAGTGTACCGACCGAGCTGCACGCCCAACTGGTGGAACTGGCCCAGG
AGTATCGACGCTGTAGCTAACGCCGCCTGCACATCCTGCTGTGGCGTGCAGGTCAACCACAAGCGGATTTACCGGCTGTA
CCGAGCCGCCGGCTTGATGGTGAGGCGGCGGAGGGTCGCGGTGGCGCGCGAATGCCTGAGCCTGTCGAGCGCACCGAACC
AGGTCTTGTCGATGGATTTCGTCTTCGACGCGCTCAGCACTGGGCGACGGATCAAATGCCTGACGGTGGTCGATGACTTC
ACCAAGGTGTCGGTCGACATCTTGGTGGAGTACGGTATCAGCGGTTTTCGTGTCACGCGGGCGCTGGACGAGATGGCGCG
GTTTCGTGGCTACCCGCAGGCGATCCGCACCGACCAGGGCCCCGAGTTCACCGGCAAGGCGCTTGATCAGTGGGCCTGTC
AGCGTGACATCAAGTTGAAGCTGATTCAGCCTGGCCAGCCCACGCAGAGCGCCTTCATCGAGTCATTCAACGGCAAGTTC
CGGGGCGAATGCCTCAATGAGCACTGCTCGCTGGTCGAAGCCAGAATCCGTATCGCGGCTTGGCGGGATTACAACGAGCA
CCGACCACACAGCGCCATTGGCAATCTCTCCCCGGCAGAGCTTGCTGCGAAGTGGCGAACCAACCAGCAGCAGCTGAAGC
GGGAAAAGTTGATATCAACCCCATAGCCTACTAACTAGGCAGCGGTACTAAAACTGGGGGCAGGCCAGTCCGAACTGATA
AAAAAACGTCCGAAACTGGCAAAACCTTATGGCCGGTTTTCCTAATATTTGCTCCGAACCCTCGGTATTTCAAGGAGCGA
ACATGCATATCCAATCGTTGGGGGCTACTGCCTCCTCGCTGAATCAGGAGCCTGTCGAAACCCCGTCGCAGGCAGCGCAT
AAGTCCGCCAGCTTGCGTCAGGAACCTTCAGGGCAAGGTCTCGGGGTTGCCCTAAAGAGCACGCCGGGAATACTTTCCGG
GAAGTTGCCGGAAAGCGTTAGCGACGTGCGTTTCAGCAGTCCCCAAGGGCAAGGGGAGTCCCGTACTCTGACTGACTCGG
CAGGGCCGCGGCAGATCACTCTGCGCCAGTTTGAGAACGGAGTCACCGAGCTACAGCTCAGTCGGCCACCATTGACCAGT
CTGGTCCTAAGCGGCGGTGGTGCCAAAGGTGCGGCATACCCGGGAGCAATGCTGGCGCTAGAAGAGAAAGGCATGCTCGA
TGGCATCCGCAGCATGTCCGGTTCGTCCGCTGGCGGCATCACCGCCGCCCTTTTGGCCTCAGGTATGAGCCCGGCGGCGT
TCAAGACCCTTTCCGACAAGATGGATCTTATTTCGCTGCTCGACAGCTCGAACAAGAAGCTGAAGCTGTTCCAACACATT
AGCAGCGAGATCGGCGCATCGCTGAAAAAGGGCTTGGGCAACAAGATCGGCGGCTTCTCTGAGTTGCTGCTCAATGTACT
CCCACGCATAGATTCGCGGGCTGAGCCCCTAGAACGCCTATTGCGCGACGAGACACGCAAGGCCGTGCTCGGACAGATCG
CTACGCATCCAGAGGTTGCACGCCAGCCGACCGTTGCCGCCATCGCCAGCAGATTGCAGTCCGGCTCCGGAGTCACCTTT
GGCGATCTAGATCGGTTGAGTGCTTACATTCCCCAGATTAAGACGCTGAACATCACAGGTACGGCCATGTTCGAGGGGCG
TCCGCAATTAGTGGTGTTCAATGCCAGCCACACACCGGATCTGGAGGTCGCCCAGGCGGCACATATCTCCGGTTCCTTCC
CAGGAGTGTTCCAGAAGGTCAGCTTGAGTGATCAGCCGTACCAGGCCGGCGTAGAGTGGACAGAATTCCAGGATGGCGGG
GTGATGATTAACGTGCCGGTCCCTGAGATGATCGACAAGAATTTTGACAGCGGGCCACTGCGGCGCAACGACAACCTGAT
CCTTGAGTTCGAGGGCGAAGCTGGGGAGGTAGCGCCCGACCGAGGTACTAGGGGCGGCGCGCTCAAGGGCTGGGTCGTCG
GGGTGCCTGCCCTGCAGGCGCGCGAAATGCTGCAGCTCGAGGGCCTGGAGGAATTGCGCGAGCAAACCGTTGTGGTGCCG
TTGAAGAGCGAGCGCGGTGATTTCAGTGGCATGCTCGGTGGCACCTTGAACTTCACCATGCCGGACGAGATCAAGGCGCA
TCTTCAGGAGCGCCTCCAGGAGCGAGTCGGTGAACATCTGGAGAAACGTCTTCAGGCTTCAGAGCGTCATACCTTCGCTT
CTCTCGACGAGGCGCTGCTGGCACTTGATGACAGTATGCTCACCAGTGTTGCTCAACAGAACCCGGAGATCACAGACGGG
GCGGTGGCTTTTCGCCAGAAGGCGCGGGATGCGTTCACCGAGCTGACTGTCGCTATCGTTAGCGCCAATGGCTTGGCGGG
TAGGCTCAAGTTGGACGAGGCTATGCGCTCCGCTCTTCAGCGACTCGATGCGCTGGCAGATACTCCGGAACGCCTAGCAT
GGTTGGCAGCTGAGTTGAACCATGCTGATAACGTTGATCATCAGCAGTTACTCGATGCCATGCGCGGGCAGACGGTGCAG
TCGCCGGTGCTCGCCGCTGCGTTAGCAGAGGCGCAGCGCCGCAAAGTGGCGGTTATTGCCGAGAACATTCGTAAGGAAGT
TATCTTCCCCTCTCTGTATCGCCCTGGCCAGCCGGATTCCAACGTAGCTCTGTTACGTCGGGCGGAGGAGCAGCTACGGC
ATGCCACCAGTCCGGCGGAAATCAATCAAGCGCTGAACGATATCGTCGACAACTACTCGGCACGAGGCTTCCTGCGTTTC
GGCAAACCCTTGAGTTCGACTACCGTTGAGATGGCTAAGGCTTGGCGGAATAAGGAGTTCACATGATTGATACATGGCTG
GCACAGTGGGGCTTGAGACTTCCCTCGAGCAACGATGCCACGTTGCGGCTGCAGAGGGACCGGAACTGGTTAT
GGAGCGCCTCGAGGGCGGTTGGCTTTTCGTCGTCGAGTTGGGACTTGTGCCTTCAGGGTTACCGCTGGGTGTGATCTTGC
AATTGTTACAAGTGAACTCTCCATTCTCATCCTTGGCACCGGTGAAACTTGCGGCGGACGATGCCGGTAGACTTGTGCTC
TGGGCTGAGGCACGTGATGGCGTTGACGATGTGGATGCACTGAACCGCTTGCACGATAGGCTGCGGGAAGGACATTCACG
ATTAGTGCCATTGCTAGAGCCCACGGGTGAGTTGGTTCCAGCTCAGATACAAACCAGCGCGTTAGTGTTCGTTTGAACGC
AAGGGGAAGGAGCGATACTGCCAATAAAATCCGCAGTTGGTAATTTCTCTAATAGATGGGCCACCCGAAGGTGGCCTATC
TGTTCATGCTGCAATGAGTTGTCGTGCCAACACCATTTCTACCGAGCCCCATCCTGGTTCAACGAGCCGAGGCCCTCGAG
GTGGGCTTTTCTGTTTCTGGAGTCCGGATATCTGATCCACAGCGCGCCAGAACGAGCGGCTGGTTTCGTGCTTTCGCCAG
TTATATCGGCACTGGCCCACATTCGTGTGCCTGCCAGTCTGTTTCCAT

FIGURE 2A

The Big Island: 84830bp (SEQ ID NO: 2)

CACCGGCGATGCATGGGATGCCGTAGGGGTCGCTGATCCAGCCGAGGCCGCCCTGGATGGTTTTCTGGTCGGCGCTGGCG
GTCTGGTTGTTGTTGCTCTGGTTGTCATTCACCTCCTCGGCCGGCGCCGGGAAGGTGCGCACGGTCCCGTCGTTGAACAC
GAACGTGAGGCTGCGGATCTGCCCACGCACGCAGGAGAGTGTCCAGTCGCCCGAGGCGGTCCCGCTGGCTACCGCGCCGG
CGACGTCCGGCAGCTCGATGCCGTTGGCGGTGAGGTTGTCCGGGCCGATGAGGATTTTGAACGGGTAAGGATCATTGACC
GTCCCGTCGACCGGCACACGACCGATCAGCGCAAACATGGCCACCGAGCCCATGAGCGTCGAGTTCTGCGGCAGCGTGTA
GGTCTTACGGACGAGCTTGCGGTTTTCCAGGTCAGATCGCTGTTGGCCGACGGGGTGCAGCCCGTCATCGATCCGCTCCA
GCGCGTTCTGTCCGCGATCGACCGCATTGCCGAAGGAGGTCGGGAAGCTGAATCCGCTCGGTTGGGTGGTGGAGCCGGCG
GCCAGCGGCTGGCCATTGGCATCAACCGCGCGGGCGTCCTGGGGCTCGATCCAGACGATATCCGATGAAGACGAGCCCGC
TCCCTGAAAGTGCTGGCCATCCTTTGGTCGCACGCCGAATCCTACGGGCAGGTCGGTGTCACCGGCCTCGGGAATGTGGG
TTAGGTTCTCGAGCCGTTTCTGTACCTGGTCCAGGATCGTCTGGCTACGGTTTTCCTGCTGTTGCTGCAGTTCCTGGGCG
GTGTTGTTGAGCTTCTGCTCGATGTTCTGATCGATGTTGCGCAGGCGGCCCTGCAGGTTCTCATTGGCGGCTTTGAGCGA
GTCATTCTCCTGGATCACCTTGCTGATCTGGTCCTTGAGCTGCCGGCTTTCCGCCACGATGGTGCGTAGTGTGTCGGCGG
GCGTGTCGCCGTCGATGCCCAGGGTTGCCGCTTCTTCCGACGTTACCGTCGGGGTTGCGGCGCCTTGTGACTGCGCACTT
TCTTTCTTGCCCATCATGCTCACGCCGATCAGGATGGCGCCGATCACGACGGGGACGACCAGCAATTTCAGGAGGGGATT
ACCGGTCAGCGCCACGACGGCCTCCTTTGGGATCGATCTGGCTGATGGAGGAGGGCAGGAGCGCGTCGGCAAGGCCGCGG
CCGCGCGTCACCAGATACACGGTAGTGGTGTCGGAGGCGTCGCCCCGGGGCCCCAAGTACGGGTGCTGGAAGGTCGCGGC
GACGAAATTGCCCATCAGGTCCCTGGGATCCAGGGCCAGGTGCTGGGCGCTGGCGTTCTGCAGCTTCACCGCCGTGATGT
AGTAGTCGTCCAGCCGCCAGGCGCCCAAGGCGGTAGCCGTGATGGGTAGGCTGGGGAGCAGGGTGGTCAGGTCGAGCTGT
CGCTTGACGCGCACCTGACCGACGCCATCCACCGGTTCCACCGTGCGAAGCGGGGCATAGAGCATCTGCGCCGCATAGCG
CGTCAGAACCACGGGGACGGGCGTTTCGCGCGGCACGGCCTTCGGTGCTTCTGCGTGCTCGGTCTGTTTCGCTGCTGCCG
ATGGCTGGGCTTCCCGGGACTGGCCATAATGCGGATCCACTGGCTCGCCGGCGACGATCCTGACCGGCTCGCGCGGCTGT
TGGTCGGCCGTTGCTTCGGTGGCGGCGATATCGATGAGCATCTGCTCGCCATTGGTCGCGTCCTGTAGGCGCAGGCGCGC
TGGAGGAATCGGCTCGTTGGCGAGCAGGTAGAGTGCGCCGCCGGTACTCTGGACGCGCAGCTTGCCCTGCAGATCCCGAG
GAACCCCAACTCGCACGTTTCTGTCGACGAAAACAATGCGTTCCTGGCCGACCGTCAATGGAATGGCCAACGGAATGCGC
TCCCAGCGCAGAATCTCCACCGCGTGGGCCAGTGTGGGTAGGGCAAGCATTAGCAAGAGCGAGCCTGTCGACTTCCGGAT
CATAGACCTCCCTCCCGCTTGGTGGGGGCGGCCGGCTCCTCAAGCTCGATACGTTGAGGCGTGTCGGAGTAGCAGTCCCA
CTGCAGGCCGAAGGGATTGGTTTCCGGGTCGACGTCGGCGCGGATAACGTGCAACGGGTAGCGGGCCAGCGCCCGCTTGA
TCTTCTCGCCGGCGTAATACTCCGTGCTGTCCATGTCCAAGTTGACGGTCCAGTCATTGATCGAGTGCTGGATCACGCGG
CCATTGCTCTCGCCAATGCCTCGACCGGGGATTTCCGAGGTGGTGCGCTCGCGACCCCTGAGCTCGCCGGCGTTACGACG
AAACTCAAAGTCTTTCTCCAGGAAGACTTTGCAGGAGGGAGTGAGGTAGGCAGCGTAGCGGAACAGGTTTCCCTTGTAGT
CCACCTCGCCGTCCTTGGGCCAACGCTGCACCTGCTGGAAGATGTAGAGGCCGAACGCATAGACGCTCTCTGGCGGAATG
TCCCACCACAACCGGGTGCTTCCTGAGCGCAAGTCGGGCGGAATGTGGATCCAGAGGTCCTTACGGGCTTCCCAGACGCA
GTAGGCCAGCACAACGATGACCATGCACAGGAAGCCGGTGATGAACCGGAAGGTGTTGATGTGTGCCTGCTGTTGCGCAG
TGTGTTTTCTGAAACTCATGGGGACTCCGTGCGTCGACTGGTCCAGGCGCCGGAGCGCGTAACCAGGCGACGGTTGCCGA
AGACGGGAAAGCGCAGCGAGAGGGCCATCTCCACCTGGCGGTAGAACCAGGTATCGGGCCGCCCCCGCTTCATCCGCCGC
AGGACGCGGCTGCCGATACCTAGGCCGAGGGCGCCGACCAGCAGCGCGCCCAGTGGAATGCAGGCAGCGTTACCTGCGAC
CAAGGCAGCCGGGATGCCAAGAACGAACCCGGCTGCTCCGCTGGTGAAGACCGTGATCCACATTTCGTCTGCGGTCAGGC
CGCCGATGACTACCGGTTGCCGGTTCAAACGGGTCGGCAGGAAGCTGAGGGTTCCATCCTGAAACAGATGTTCTTCGGGC
ATGGGAAGGCCCTTACATGATGGCGGTGGCTTTGGTGACGAGATAAATGATCAAGATCAGCAGGCCGACACCTACGGCTA
CGCCCGCTCCGAGATCCGACCACTTCTTCTTCCCGTCATGGATGGCGTGATAGGTGCCGTAGGTATGCCAAGCGACCCCC
AGAAAGACAGCCGCGCAGATGAGCAGCGCGAGGAGCATCGCTCCGTCATAGCCGAAGTTCTGGATGGTTTGCATGATGCC
CGATCCCTCCCCACGGCTAGGTGCCTCGGGTTTGGGGAGTGCAGCCAAGGCGATACCGGGGAGCGCCAAAGTGATGGCGG
CCAGGCTCTGGCAGAGGGCGGACAGTTTCTGGAGGGTGAGTTTCAGCATGAGTGTCTCCGTATGGGTCAGCTGAGGAGGA
AGAAAGAGGTGATGCCGAGGAGGACCAAGATCCGGGATCGCGGAAGCGCCACCGGTGGCGCTGTCGAAGGTTGTTGGTGG
CCCAGCCGCGCCAGGTGCTGTACATCGCCCAGGCGGACCACAGCAGAAGGAAGGTCATCGCTGCTCCAATGAACAGTCCC
TCGCCGGCCGATGGGGGAAAGCCAGCGGCGGCCTGGAACGCTGATGTCTGGGCTCCGCTCATGCTCATGGCATCGGCCCT
CCGCGCAGGGTGTAATTCCCTGATAAGTCGGAAGGGTCACGCGGTTGGGCGCGGCTGGGCGTCATGTAGTCCTGCAGGCC
TTGGCGGATGCGCTGCAGGTCAGCGGCCAGGCGCGGGTAGTCGAAGTAGAAGCGCTGTCCGGGCTCATCGGCGCCCTGGG
CACTGCGGCGGGCGGTGTCCTCGAGGGCGTTGAGCTGCCGGATCATCACCTCCAGGTTCGCCTGCTCGGAGGCGCTCGCT
GCGAAAGCTCCCTGTGCGACCAGCGCACTGCAGGCCAAGACGGCCAGGCAGCTTCGGCGAAGGGATGGATGTGGAGTGTT
CACGGCGAATCTCGTGCAGATCAGGGTGCGATCAGCTTGCCGATCGCCCTGATTCCATTCCGCAAGAAATACTCGGCGCG
GGTCCGGGGTATTTTTTTATGAGGTTGCGCGCTTGAAAAACAGCAGGTAGCTGACGGGCAGAGCGACGCCCCAAGCCCAG
CCTTGGGCGTTTTGATTACCGATCTCGGTTACGCCCCGCATGAGTGGCTGAACGCTCACCAGCTCCCAGCCTTCCTGGCC
CAGACGCACGACCTCGGCCCAGCGCAGAGGGTTGAGCAAGGCGTGGTCGTCGGGTTCCAAGGTGTCTTCTTTGTGTTTGA
ACATCGCGACGCCCTTGGTGAGGGTTTTGTAGGGAATGACGATGTTCTCGACATGGTATTCGAAGGGCTGTTTTATGCTC
ATGGTGTGCTCCTGGATCGGTTCGAGTAGAACCGCACGGTTTCAAGAGCAGGAGGAGGGCCTCAACCGAAAATACTAGGC
CTAGCACCAGTGAATTTCTGGTGGTTGGCGGGTTTAGAGACACGTATGGACGTCTGGGTGTATATTTCATATATATTTCT
AAGGGAGACGGCTGATGCTTAGAAACATCTCTATTGGAGTTTTGCTAGCCATGGCTGCTATGTTGGGCAGTTATGGGGTG
GCTGCCGCTACATTACGATGCGGGTCGGCAATTGTTAGTGAGGGCGACTTGATTGATGATGTGCTTAGAAAGTGCGGCAA
CCCTGATAGCCGTAAAATTGAAGGGCCCGCAGTGGATGGTAGTGGCTATATAGTGCGGGGGGCTGCTACTGTCGAAAACT
GGGTATATGGACCAAGGAATGGATGGTACCAGAAGCTTAGGTTTGTCGATGGAAGACTAGTTCAGATAAAAGGCAGTATG
GACTAGGGTATAGCCGTGGATGGTGTGTTTTCATCCACGGCTATAAGTCTCATCCGGCAGATGATATAAGGGTAAGGATA
TTTGCGATTGGTAGGCCTTGTGCGTCGGAAATAAACACGGTTGTACTGGCGCCTGCACGGGAAAACTATCTGTTGTAGGT
TGTTCGGATTAGACATGCCACCGTTGTAATTGGCTTGGAATTGCTTGCTGGAACTTGCTATGTCTAGAAGTTCAAGTAGA
GTGCTTTGGTTGGCGGATTGTGCGGAGTTCATGCTGATTGATGTGTTTGATCCTGTCGCTCGGAGCGCCTGCTGGATGTT
GTTTGAAGTGAGTTGGGCACCATTAAAGTTGGCGCTGAATTTTTTAGTGTTGCCCGCGATGTTGATGGCCTGTAAAAGGG

Figure 2

FIGURE 2B

TGGATATATTCGCCGCCTGAGCGGTGTTGACGCTGATATTGGTACCCGAGGCCGCAGCGCTGATCGCTTGCTGAAGGTTG
TCTGCTGTTAGTTGTGCACCATTGAAGTCGGCGCTAAACTTTTTCGTGTCTTTGGCAGAGTTGATGAGGGCTAGTAGGGT
GCTTATATTAACCGCCTGTGCGGTATTGACGCTAATGGATGTTCGGGTTCCTGCGGCGTCTAAAGCTTGTTGAATGTTGT
TTGAAGTAAGTTGAGCGCCATTGAACTCTGCGCTGAATGTTTTTGTGTCACCTGCTGCATGAATAGTTTGAAGAAGGGCG
GTTATATTTGCCGCTTGTGCGGTATTGACGCTTATGCTGGTGTTTGTTCCTGCCGCATTTACTGCTCTAAGTAGGTTGTC
GCTGCTGAGTTGGGCTCCATTGAATTCTGCGCTAAAATTCTTTGAGTTGCCTGCTGTAGATATTGCGGAAAGAAGTGTGT
TTATATTAATGGCTTGTGCGGCTATTGTGGATACTCTGGTTAGTGATCCTGCTGAGTTTATGACTTGCTGAAGGTTGTTA
GCAGTCAGCTCTGCGCCGTTGAAAGTTGTTGAGAGGTTTTTTGTGTTTTCTGCAGTGGTTGTTGCCTTGAGAAGAGTTTC
TATATTTAATGCTTGCGCGTTGCTTATGCTGATGCTTGTATTTTTTCCTGCTTTTGATAGGGCCTCTAGCATGTTGCTTG
GATTTAGAAGGCGTCCGTCGAAATCAAGGATAAGTTTCATCTCAATGCTCCGCTGTGACCTAACGTTAAAGTCTTGGTGT
GAGTACTTCCTAGGCTGAATGAGTTGGCTACAAATCTCTTTGCCTTATGTCAATATCAGGCATTCGCTGTGACCAATATC
ATTTTTTGAGTCATTTTTTACAGCTAGCTATCTGCTCCTTCAGTTTTCTGTCGTCCACGTACCGCTGCATGATCGACAGC
ACCAGGCTGTTGATGGCAAGGATCTTCATCGCGCGGGTCACTGCGACGTAGATCAGGTTCAACTCATCGTCGCGCTTGCC
TGGGTCGGTGTCGGGGGCCAGCGGGTCCGCGTTGAAGTCGTCGTACAGGCAAACGAAATCCCATTCCAGCCCCTTGGCCT
TGTGTGCGGTGGTCAGGGTGATTGTTGCGTCCAGCTCATCGTCAAGGGTCAGTGAGCGAAGCTCAAGGATCCGCGCAGGC
AGATCAGGGTAGGTCGATATGATCTTGATCGAGCGAAGCATCTCACCGTCCTGGCTGATCTCGGCGATCTCCACGTACTG
GGTGTAGTCGCGGTAGTCACGGAGCAGTTTCTTGTTCTGGACGTTTTGGCGCAGGCCTCGGCTGAATGCGTACAGATCCT
CCAGGTCGCGCAGCGAGTAACTGTCGATACCGCCTACCCAGTGGAATTTGGGCTCCGGATGATTGCGGACCAGCTGCAGG
GCATTCTCGATGACGCCGATAACGGTGCGGTGAATGAAAGTGCGGTGAGGAAGGTCCGGCGGGAGGGACTTTTTCACCAG
CGTCTGCGGACCCAGTCCTTGAAGTTTCCGTGTTTCGCCCTTGTAGGAGAGGATGATGTTGGCCACGTGTGCGATCGCGG
GGCCGAATCGCCAGCTCTGGGTCAGGTAGTGCTCCTCGGCGCCGGCCATCCAGTCGCTGTTCAGGGCATCTTCTGCGCCC
CTGAACCGGTAGAGCTGCTGATGGGGATCGCCGACGATAGCCATTCTGATGCGCTGCCAATGGGCAATGTCCGCGATCAC
TGGGTTGATGTCCTGCCCCTCGTCCAGGAGCATGCAGTCGAAGCGCTGGCTCAAATCGGGCTTGCTCAGTTGATACAGCT
TCAGGTAGCCGTCAAGGGGCATCAGCATGCCGGTGTCCTGGAGATCGACCATGCGCCTCCAGACTACTCGCGCCATGTCC
AGGCCCTGCTTGAGGAAGCGTTCCTGAGCACTGGTGAGGAACGCCTTGTCGCGGAAGCGCGGGAAGTGCGGTCGGCCGAG
TTCCGCGTCGGCGCTGGCCATGTAGTTGTTCAGCGTGGCCAGCACGTCACGTACCAACTCCCAGTCTTGGGTATCGAGTC
CGCGGGCGATATCGGTCAGTCGCAGGTTCTTCGTCTTCTTGTGGGCGTACTGAATGCCGTACACCGCATGAGCCAGACTG
TGGGCGGTCTTGCACACTACGTTGCGGGGAAACTTGCCCTTCGCGGCTTTCTCCACCGAGCTGTTGTAGCAGAGATAGAG
GATTCTCAGGGTAGGGTTCGACCTGGCAAAGCCCACCAGGGTGGTAGTTTTGCCAGTGCCTGCGAAGGCTCGCACCAGGA
TCTTCGGTGCCTTCGACTGGATGATCGGTGACTGTTCGTGAGTCCACTGCACGGCGGGTCTCCTACAGGTACTTTTTGAA
GGTGGATGCTGTGATGCATACCGCTACGCCGAGCAGCGCGGCGCAGGGCAACAGGATGAGCAGGGGATTGATGCTGATGG
GGATTGCCAGGTAGAGCGTCCAAGGGACGACCGCTAGCGGAATGATGCTGCCGCGCGCCTTGTGGTAGAGGTAGCTGGAC
TCCCGGCCGGCGCCGAACTTGCGCAGGTCCCGGCGCACCAGGCCGTCCACCAGGCCGGTGAACGCGGCCATCACGAATAA
CGGGATCGTCATGACCAGGATCACCAGGCGCACGACGAATGTCAGCACCGTGTACAGCGCGGCCAGGCCGTAGTTCTGCA
GCGTGGAGACACCGTGGGCGGTGAGATAGCGAACGTCCAGCGGGCTCCGTGGCCGGGCCTGCGCGATGGTAGTCATGTTG
GTCATCCAGTCGACCATCCCGGTCTTCACGAACAACCAGTCATAGGCCAACTGGGCCAGCCAGGTGGCGGTTCGACCTGG
CTCCTGCACGACGACGCTGTGTAGCAGCCCCTGCGACAGCCAACTGAGTTCGTACTCAAACATGGCCTGGGCATGCTTCC
AGCCCGCGTCAGGCCAGAAGAAATACAGGCATACCCACTCCACGATGATTGCGCCGATCAGCGACCCGAACATCACGCCG
ATTCAGGAGCAGGAGTAGGGGGGCAATCGAAAATGCCAAGACTGGTGCCGTGGGTTTTCTGGGGTAGCCGCGGGAGAATG
AACAAAGATGTTCAAACGTTCCTATAGGAAACTGTGTTGTGAACTCCATAGAACTCCCCTTTTACACGATGCAGGTTGTT
GCTTGGAGTGTGTATCCACTAGTTCTCGGGGAGGCCGTCACTGTAGATTTTTTGGGGGGGTATCTCGGCATAGGGGGGAT
TGAATAGCGCTGTGCCTGCAGTCCTGAGACCTCATTAGGAATCATCTGATAGAGGGAATTTTTCCATGCCGTCAGTCTTT
CGGCATGAAGTTGAAGAATTTCTTACAGCCTTTTGATAGCGGTTTCTCCACTCCGAGTGCTGCGCTCAAGCTGCTCCGCA
TGCTCGGTGGCGCCTTGATGTTGTGCGTGCTATGCAGCCTGATATTCAGTGTGAGCATGGTTTTAAACCATCAGGTGTCC
CTCAGTCGGCAAGCTATGAATGTGGCTATGTACGAAGCGCAGCTTTATTTCGAGCAGCGCGAGGCGTTGCTCAATCACTT
GAGCGGCAATGTCGTGCCCTTGGCCGCGGGTAGAGCGCTCGTCAACGAAGCGCCGAACAATGTGAGCATCCTGCCGTTGA
GTGACGGAGGGCGAGGTCTGCTATTGACCGCTCGCACGCTCGGTGATCTCCGGGAAAAGCGGCTGGCACTGATGTATCTG
GTCGATACCGACAAAGGCCCTCTGGTTTACCGGCTTACCGCCGATGGTAGGCCCTCGGCAGCGATATCCAGCACGATAAC
CAAAGAGGTGTACCGAGCCTTGCTGGCGACTCCGTCGGCGCCTGTTCACTGGGTGACTGACGGTGGTACCCCTCAACGGC
TGTACCTTTTTGAATCCTTAGGCGATGAGCCGGGCGAGGGGTGGCTAGGCCTGGAGATTCTCGGCGAAGACCTCGATTCG
ATGTTGCGCCGGAATGATGCCGGAAACTACATGCTGCTGGATCAGCATGGGCAGGTCGTACTCGCTACGGACGCAGAGGC
GCTGGGGAGCGGTGCGTCGCGGACGCTTTTGCGTGGAGACGGCTTCGGTTTCATCGGTGCTGGCCCACTGCCGCAGCATA
TGGTGCTTTTCCAGCACGTGGGGTCTTCGAGCTGGGATCTGATCTATCACATCGGTATCGGTCGCCTGTTGCTGGCTCTG
TGGCTCCCTCTGTTACTTGCCTCTGCGTTGGCACTCGCAGTCGGCATCCTACTGCATTGGCTGGTGCGGAGCATCGAGCG
ACGCTTGATAGAGCCCGCAAAGCGACGCCTTGAAGCATTGAAGGAGAGCGAAGCCTTTTCCCGTGCAGTTATCCAGGCCG
CGCCCGTCGCGCTGTGCGTGCTGCGTCGTGCCGACGCCGCAGTGGTCCTGGAAAATCCCCAGGCGCGCCAATGGCTGGGT
GATAGCGAGGCGATTGCCCACGACGCGCCGAGATGGATTTCCCAGGCGTTCGCAGGAGGTGTGAAGTGTTCTGGAGAAGA
ACTGGAAACCGAGGCAGGGCTACATCTTCATCTCAATTACACGCCCACCCGCTATAACGGTGAAGACGTATTGTTCTGCG
CCTTCAGTGAAATCAGTGCACGCAAGCGGATGGAGGCGGAACTGGCTCGCGCAAAATCCCTGGCGGATGCTGCCAATGAA
GCCAAGACGCTGTTTCTCGCCACCATGAGCCATGAAATCCGCACACCTCTGTACGGCATGCTTGGCACGCTTGAGCTGCT
TGGGCGTACCGAGCTGAGTCGGCAGCAGGCCGGTTACCTAAAGGCAATCCAGCATTCCTCGTCGACCCTGCTGCAACTGA
TCAGCGATGTGCTTGACGTATCCAAGATAGAGGCCGGCCAACTGGACCTAGAGTGCGTGGAATTCTCCCCGCTGGAATTG
ACCGAAGAGGTCGTGCAGTCGTTCACCGGTGCCGCGCAGGCCAAGGGGCTGCAGTTGTATACCTGCCTCTCTGCGGAGCT
GCCGCTGCGCATGCGGGGGGCCGCGGCGTCGATCCGGCAGATTCTCAACAACCTGCTGAGCAACGCGGTGAAGTTCACCG
ACAATGGCTATGTCAACGTCCACCTGAAGGCCAGCGTGGTCGATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGAT
ACCGGCATGGGGATCAACGTCGAGGATCAGCCGCGTCTGTTCGAACCGTTCTACCAGATACGCCGCTCCGAGCATCCGGT
CGCAGGCACGGGCCTCGGCTTGTCGATCAGCCAGCGCCTGGCGCAGCTAATGAATGGCAGTCTGAAACTGGTCAGTGAGC
TGGGGTTGGGCAGCAGCTTTAGCCTCAGGCTTCCGCTTGAGCGGATCGCGATGCAGGCTGAGCCGCAGGACCTAGCCGGG
TGCGCCGTCCAAGTGCTGGCGCCTGTCCGCGACCTAACGGAATGCCTGTGTGGCTGGATCTCCCGCTGGGGTGGAAGGGC
CATGGTCGCGACGCCGAGGTCGCTGGACGAGGCGGACGCGACCTCGCTGCTGGTCAAAGTGTTACTGCTGGAGGGGGCGC
CGATGTTCGAAGCATGGCCAGGATGCCGGGTGGAGCTTTCCCCTCAGGGTGATATGGAGCCGCAGGCACAGGGCCGCGAC

Figure 2

FIGURE 2C

TGGCTGCTCGGGCTCAACAACCTGAACGGCCTGCATCGTGCTCTGGGCCTGGCCCATGGGCGTCTCGCTGATCCTTCGAC
GCCGCCGATACGGCTGGCTCCGTTGCGCAATCTAGGTCTCCGCGTCCTAGTGGTGGAGGATAACGCGATCAACCAGTTGA
TCTTGAGGGACCAGATGGAAGCGCTGGGCTGCAGCGTGGAGCTGCTCTTCGATGGTCGCGAGGCGTTGCTGCACTGCCAG
ACGGCCTGCTTCGACGTGGTGCTCACCGATATCAACATGCCGAACATGAACGGATACGAGCTAACCGCGGAGCTACGGCG
CCAAGGGTTCCGGCAGCCGATCATCGGCGCGACGGTGAACGCCATGCGTGAGGAGCGCGAGCGCTGCATGTCCGCCGGGA
TGAACGATTGCCTGGTCAAACCGGTGGATCTGAATGCCCTTCAGAACTGCTTGATTAATATTCTCAAGGTGGATCGATGA
GCTGGAAATCCTATCGGGTGCTGGTGGTCGAAGATCAGCCGTTTCAGCGCGAATACCTGCTCAACCTGTTTCGCGAGCGC
GGCGTGCAGTACCTGGTAGGTGCCGGCGACGGCGCGGAGGCGTTGCGCTGCCTGAAGCAGGACAGGTTCGACCTGATCCT
CAGCGATCTGATGATGCCGGGCATGGATGGTATCCAAATGATCCTGCAACTGCCGTATCTCAAGCATCGTCCGAAGCTGG
CGCTGATGAGCTCCTCGTCGCAGCGGATGATGCTCAGTGCCAGCCGGGTCGCCCAGAGTCTCGGCTTGTCGGTAATCGAC
CTGTTGCCCAAGCCGACTCTGCCCAAGGCCATCGGCCAACTTCTGGAACACCTGGAAAGATGCCTCAGGCAGAAGCTGGA
GCCGGAAACCGACGAGACTCCGCATGGGCGCACGGCGTTGCTGGATGCCCTGCATAACGAGCAACTGGTGACCTGGTTCC
AGGCTAAGAAATCCCTCCACACCGGGCGCATAGTCGGCGCCGAGGCGTTGATACGCTGGAGCCACCCGCAGCATGGCCTG
TTGCTGCCCAGCTGTTTCATGAGTGATGTCGACGCTACCGGTCTGCACGAGGCGTTGCTCTGGCGCGTGCTCGAACAGAC
CCTGAACGCCCAGGAATCGTGGCGCAGGGCGGGTTACGAGATTCCGGTTTCGGTGAATCTGCCGCCGCACCTGCTCGATA
ACCAGGAACTTCCGGATCGACTCTATGAGTACGTCGGCGCTCGCGGGGCTTGTACCAGCTCACTATGTTTCGAGTTGACC
GAGAGCAGTGTCACAACTCTGTCAAGTAACTACTATGCAGGTGCCTGTCGCTTGCGCATGAAAGGGTTCGGATTGGCCCA
GGACGACTTTGGCCAGGGTTACAGCTCGTTCTATAACCTGGTCACGACGCCTTTCACGGAGCTGAAGATCGACCGCTCCC
TAGTCCAGGGATGCGTAGAGGATAACGGCCTCAATGCAGCTGTCATCAGTTGTATTGAGTTGGGTCACCGCCTGAATCTC
GACGTGGTGGCCGAAGGCGTGGAGACCTGCGAGGAACTGAATCTTCTTCGTCGTCTTGGCTGCGACCGGGCGCAGGGTTT
CCTGATTTCTAAGGCAGTGTCTGCTCGTGAGTTCGAGCGGCAGTTAAGGGAGGACGGCCCCAGCCTCCTTGTTTAACGCA
GTATCCCCATTATCGCGGAGTCGATCGCAGCCAACCACCGTCAGCGCAACAGTGTCGCTGAAGGAGGCCGCATCCCGTGA
AGTCTGCTAGTGCCTTGGAGCACGACAACAAACTTTTGCTCAAATGGACAACCCTCTCGCAGAGCCTGAGCATCGGCTTG
ATCTGTGTGGTGGTGCTGACCGTATTGCTGTTCAGCATCTGTTACTGGTCGCTGGGGAGATTGTTTCAGGAGGAGGAGGA
CAAAGTCTCCTTCCACTTCACCCGTATGATGGATGTTATACGGGAGCATGAGGTATTTCTTGGGCGCATCGCTCGAAAAA
GCGACAAGACCACCCAGAAGTACGACTATGACGTGGTGCCTTTGCAGCGGCACTTGTTGGCAAAGGAAAACGGATTAGCG
GTCTATGAGGGACGGGAGTTTTCCTTTGCTATGCCATTTCTACTGGCTACCAAGCACGCGTTGAGCGCCGATTCCTCGGG
AGATCCGTTTTCGCTCGGTGTATTGCTCGCCAATTTCTACGGAAGCTTCTGGAGTGTTTCCGCCTATCCCGCGCCACAGT
TACTGATCTTTGATCTTTCCGGCAGCACCCGCCTGGCAGTGCCGTCGATTCCCTCCACAGCGCAGCGTGACAGGTTGAGC
GGAAGCTATCCGATGATAGTCGAGCGCATTCTGGCGCGCTTGCGCACCCGGCCGGTGGGGGAGGACGCTCAGCGTGTCCA
TTGGATACGCGCTGATCGCTATCGCGACTCGGCGCTGGAGATGTTGGGAGTCGCCCGGGTTGATCTGCCGGAAACACTCT
GGTGGCACGACGAGCCGAACCATCTGATCATCGCTGCGAGCCTGCTTGATCTCAGGCGAATCAATGACTTCGAACAGTTG
GTTGAGCGCCCGGCATTCGATTCGTACAGCCTGGTATCGCCGGATGGCGAGGTATTGCTCGGCGCGGCCCCTGCGACCGG
CCTGAGGGATGGCCTGAACCTCACCCGACAGGGGGTCGCCGTTCAACTGCTCAGCCAGCCTGAGAACGGTTGGCTCGCGG
TCTACCGAACCGACTACGGCAATTTCTTTCGCCACTCCCGGTGGCTGGTGGCAGGTCTGCTGCTGACCCCGGCGCTGCTC
CTGGCCGGTTGGCTCGGGATGCGTTGGTACACCAGCAGCGTCGTCAACCCGGTGCATCGGGCGCACCGGCAACTGGTGGA
GAGCGACACCTTCAGCCGGACGCTGATACAGACCGCGCCGGTGGCTCTGGTGGTGCTGACCCAGGATGACCAGCAACTGG
TGACCTGCAACCACTTGGCCGCCCAGTGGCTGGGCGGGCCCACGGAGATCCTTGGGCTGACTTCCAACTGGAAGCTTTTC
GATGCGCGTGGGCAGGTACCAGGAGACATCTGTATCCAGGTCGGTGGGCGCTATTTGCAGACCGCCTTCGCGGCGACCCG
CTATGCCGGCACCGAGGCGGTACTGTGCGTATTCAACGACATCACGGTCCACTGCGAGGCGGAGACCGCGCTGTCCAATG
CGAAGCGAGCAGCGGATGCCGCCAGCCAGGCCAAGACCCTGTTCCTGGCCCGCATGAGCCATGAAATCCGTACTCCCCTG
TACGGTGTCCTTGGCACCCTGGAGTTGCTCGACCTGACCACCCTGAACGAGCGGCAACGCGCCTACCTACGCACCATCCA
GAGTTCGTCTGCGACGCTCATGCAACTGATTAGCGATGTGCTGGATGTCTCGAAGATCGAAGCGGGGCAGATGGCTCTGA
CCCTGGCCGCCTTCAATCCGCTGGACCTAGTGCGGGAAGTGCTTGGCAACTTTGCCGCCAGCGCCATGGCCAAGGACCTG
CAGTTCTATGCCTGCATCGACACCGAAGTGCCGGCGCAACTGATCGGTGACGTGACGCGGATTCGCCAGGTGCTCAATAA
CTTGGTGAATAACGCGCTGAAGTTCACCGATATCGGACGGGTGGTCCTGCGCGTGAAGTTGCTCTCCCGCAATGATGGTC
GAGCCCTGTTGCAGTGGCAGGTCGCCGACACCGGTATCGGTATCGCACACGAACAGCAGGAGCGCTTGTTCGAGGCGTTC
TACCAGGTTTCGGGAGCGCACCATGCCGGCGGCACGGGGCTAGGACTGTCGATCTGCTGGCATCTGGCGGAAATGATGGG
CGGTCACCTGCGAATGGTCAGCGAGACAGGGCTCGGCAGCAGCTTCAGCCTGGTGCTCGAGTTGCCCGAGGACGAACAGT
CCGGGCTGGCTTGCCGGCCGGGGCTCTTGAAATCCGCTTGCGTCCATGTGCGCTCGCCCGTGCGGGAGCTAGCCGACAGC
GTAGGGGCGTGGCTGAAAGCCTGGGGCTGCAAGGTCAGCAGCGGCGAGGCGGCGCCCTCCGAGCTGGAGACTTGTGTGCT
TCTGGAGCTGCTGCCGATGGCGGCCGGGCCTGCTTCTTCGCCCTGGCCAGGCCCCCGGGTGCGCGCGTCCATGGATGCGC
CTTGCCAGCCGGAGCTGCGTGAGGACGGCTGGCGTGTCGGCCTGCACAACCTGGCGGGAATCGGCCAGGCCCTGGCGCAG
GCTCTGGGTGGCGATATCCCCGAGCAAACGCCGGCAAATGCCTGCGCCCGCTCGGGGAGACTCGACCTGGAAGTGCTGGT
CGCCGAGGACAACCCAGTCAACCAGGCGCTGCTTCGCGAGCAACTGGAAGAGCTGGGTTGTCGCGTGAGCCTTGCCGGCG
ATGGGCGGCAGGCCCTGCAGCTGTTCGACAGTGGTCGCTTCGACCTCCTGCTCAGCGACGTCAACATGCCGAACATGACC
GGCTACGAACTGACCCAGGCGCTGCGCGAACGAGGCGAGACGCTGCCGATCATCGGCGTGACCGCCAACGCCCTGCGAGA
AGAGGGCGAGCGCTGCCGGGCAGTGGGAATGAACAGTTGGCTGGTGAAGCCGATCACTCTGCATACCTTGCATGAACTGC
TCAGTGAGTTCGCTCGCGCAGGTGTCGTGCTTCCCGCGCAAGCGCGAGACCTCGGCCCGCCCGCGCAGCTCGACGACGGT
CTCTCACCGCAGGTGCCGGAACGCATGCGCGCGCTTTTCCTTGAGACCATGGGCAAGGACCTGGAGGCCGCCCGGCAAGC
GATTCGCCGCAACGACCCGAAGGGGCTGCAGCAGGACCTGCATCGCATGGCCGGCTCCCTGGCGGTGATGCGTGCGCGAA
CGCTGGTGGTGATGTGTCAGGGCGCCGAGGAAGGCCTGCTGGAGTCGCGCCTTGAATGTTCCGCCGTGGAGATTGGCGAG
GTGCTCGTTCATATCGAGCAGGCGCTGGAGTTTGTGAGAAAGACGGGCTGATGCGTCCGGGGTCAATAGTTGGAATTAGA
ACACAAGAGAAGCCTATGAGTAAGCTCAAGATAGTACTGGCCGATGACCATCCGATCGTGCGTATGGGCGTATGCGACAT
GCTCGAGCGCGACGGTCGGTTCGAGGTGGTGGGCGAGGCCTCCACGCCCAGCGAACTGGTCGAGGTGTGCCGGCAGAGCG
AGCCGCATATCGCCATTACCGACTACAGCATGCCCGGGGACGAGCGCTACGGCGATGGACTGAAACTGATCGACTACCTG
TTGCGCAACTTTCCTCGTACTAAGGTGCTCATCTTCACCATGGTCGGCAACCGCCTGATCCTCGACAGCCTCTACGATCA
CGGGGTGTCCGGCGTGGTGCTGAAGAGCGGCGAACTCGACGAGCTGCTCTTGGCGCTCGACGTGGTGAAGCAGAACCGCG
TCTACCGGGGCGCGAACATGCTCGACCCGACCAGTGTTCTGGCGAACCGCGACGAAGTGGAAAGCCGCTTCGCGCGCTTG
TCGATGAAGGAGTTCGAAGTACTCCGTCACTTCGTTTCCGGCAGCAACGTCTGCGATATCGCACGGCTGCTGAAACGTAG
CGTCAAGACCGTAAGCACGCAGAAGGTCTCGGCGATGCGCAAGCTGGAAGTGAACAGCGACCAGGCCTTGATGACCTTCT

FIGURE 2D

GCGTGCATGCCAACTTGTTCCATTGACAAGTTTCCACCGCGGAGCGTTCGAGACGCGCCGGCGGATGTCTCAGTGAAGCA
GGCTCCGTCCCTGCTGGAGATTGCCGCCGTCATCGATCCAGCTGAAGCGCACCTCGGCCTGGCCGCCGGGCAGCGCCTGG
CGGAGCGCCAGGACGCGCCTGGAGAAGGGGGGCAACAGTTCGGCCGCCGCCAGGCGGGTGTGGCCGCCAAGCTCGGCGCC
GGCGAGGCTGATGAAATAGGGAGTAGGGTTCTCCATCTGTAGTGTTTCGCCCCGCCGCTCGAACCTGAGCTGGTGTGGCG
CTTCTTCCGCACCTCTGGCCAGAGGGCGAGGGCGGTAGATGACTTTCATCTGGGTGCGCAGTGTAACGGTCAGCCGTGAC
TGTCCTTCGGCCAGCAGCCCGGTGGCTTGCGGCGGGATTTCGTAGAGGTTCAACCAGTACAGCGATTCGCGGTCCCCTGG
CAGCGCCTGGCCGGTCCGCAGCAGGCGCAGGCTGCGTTGCTGTCCGGGTTCGAGGCGGAACACCGGCGGTAGCGGCATGA
TCGGCGCGCGGGCAGACTGCGGTGTCGAGTTCGGGGCGCCATCGTCGATCCAGGTCTGCACTATGACCGGATAGGCATTC
TGGTTGACCAGCAGCAGCGACGTCTCGCGGTGCCCCTCGGCGATTATTGCCCGAGTGCGCTCGGCGGTGACTCCGGCTCG
GGCGTCCAGCACGGCCAGCATTGCTAGGCCGAGCAGTGCCCCTTGCAGCAGGATCTTACTGGACACGAATCACTACCTGC
GCGCTGGCCTCCACCCGTCCCGGGGTAACGCTGGGCATGCTGCCAACGGTGAGTTTTTCCAGGCGCGCACGGAAGGTCTC
GCTGTACTGGCTGATGCCTCCCGCTTCGCCAGTCTTGTTCGAGGCGTTGCCGATCACTGGATACCAGCCCCGTGTTTCGG
CATTGCTGCCCATCGCGGAATCCTCGCTTACCAGCAGGTTCATCGCACTGCCGTCGCGCAGCAAGCGGATACCCACGCCT
TGGGCCATACCAGGCTGGCCGTAGCGGTCGGAAACTAGGTAGGAGAGCCCGCCATTGGCGTTCCTCAGGCCCAGTGCAGC
GGACGCGGCCTGCGCCCCGGCTGAAGCCCTGATACCCAGTGCAGTACCGCTGGAGGCGGTGCTGTTGATCACTCCCGTTT
GGCACTTGAAGGCCACCTCGAAGGGCATCTCACGGTTCGCGCCGGCATTAATCTCACTGAGGGAAATCGACGGGAACACA
ACGTGGGGCGTGACGTTGGTTACGGAACAGGTGGGATAGCGCTTGAGCGTCACGTCGTTGTACAGGCCGATGGCGCCCGG
CCAGTTGTAGTGCCAGCCGAGGTAGTTGGCGTTATGGCTGGCGCCGACGTTGGGGTAGGCCAGGCCGGGACCCTTGATGG
CGATGTAGCCAGCGGGCTGGGTGTAGGCGTACAACCGCGAGGCGGTAGTCGGCGAGTAGTAGCGGATGAACTCCAGCGGA
GCGCGGAACAGTTCGGCGCGTACCGCGCTGAGGTTCTTCGCCTTGACCAGTTGAAAGCCTCGCGAGTCGATATCGAGCCC
GCCGAGCAGACGCTCGCGCCAGACATCGGTGAAATACTGCCCGGTTTCCACGTGGGTTAGGCGCAGCAGCACGTTGGGCC
AGGCGGTGCGATAGGCGGACTGCAGGCCAATCGAGTTGCCCGCACTGTCGCCTCCTAGGTACCAGCCGCTGTAGAGATCG
TCGGCATTGGTGGAGAACATCTCGTACACATCGTCCTCTGGAGCACAGCGGAAGAACACGCGCTCTGGGTCGTAGCCGGC
GCTGTCGCCGTATTGGCTGAGCGGCGCCACCGAGCTGGCGATCAAGCTACCGTCGGGCTGGAAGCTGGCGTCGCTGACGT
TGATCACGCTCGGTAGACCGAGGGAACCGTTGCAGGTATCGCAGGCTCCGGCCCAGGTGCCGGCAGAGCCTTCACCGGGA
CGTATCTGGGTGTTGGAGGTAGTCGTGGCGTTGCCTACCGCCGTGACCTTGTAGCAGGTCGCCCAGCTAGACGAGGCGAG
CAGCGCGAGGAGCAGAGGCAGGACGCGCAGGGATGTTTTCATGGTGCCTCCGAGACGGGCGTGCAGGTGCCCTGCAGGCG
GATCACCGCCTGTCCGGGTTCGATAGCCTGGCGAGGGCCAGTGTCGAGGTCGTAATCCAGCAGACAGGCGTCCCCTGGGC
GTTCGCTCCATTGCACGCGCAGGCGGCCCTTGTCGCCCTCGGCGCGGGCATAGACCTGGCCGCCTTGACCGACCATGCCG
ATGTTCACACCCTGGCTGTCGAGCACATTGGCCCCTAGCGGTAGCGCGCTGCCGTCGGCGAGTTGGGCCTGGATTAGCAA
TGGGTGACCGGTCAGTGTGCGGAACTTCACATGCACGGCGGCGCCGGCGTATGGCGCGATCTTGCGCTCGGTCTCCAGCA
GCTCGGCCTCTTCGCCCATGCCCTGCGGATCGAGGCTGACCGGGTTGTAGCGGTAGGGCGAGAGTGATGGCACCACGGCG
TAGCCATTGCCGTTGACGCGCGCGCCCTGACCACCGCGTACGCCAGCTCCGCTGGCGCCCTTGGCCTCGACCAGGGCGAA
AGTGTCGCCGACGTAGGGGCCGAGGGTCAGGCCGCCGCGATGCAGCACGGCGGCGCCGCGCAGGCCGGTGTTGAGCTGCC
GGTAGCTGCTGTTCTCGGCATAGCCGGCGTTCAGCGTCGCCACCGAGGTCTGTTTCTGCAGACTGCCGTTGAAATCGCTG
CCGTGGTTGTCGCTGTCGCGCCCGGCGGCAATCGCGTAGCTCAGGCTGCGGTCTTCGTCGAGGGTGCCGTTGAGGCCCGT
CTGGTAGCTGCTGCCGCGGCTGTCGCCGGAACGCCGGGAGGCCATCGCGCTGAGATTCGGGGCCTGGCTGGAGGACCCCA
GCGGCATGGAAAGTGTCAGGGTTAAGGTGTTGCTACGTTCGCTGCCGTGCCGCAGGTTGTATTGCGGCGGCAGTGACGGG
TCGTAGTCGTCGTTCAGATCGCGGTACCAGGTGGTCTGCTGGCGCGAATAGGCGAGGTTGTAGCTGAGCTGGCGCCAAGT
GTTGCTGTAGCCGAACTGCAACTGGGTGTCGCGGCTCTTGCCGTCGTAGTAGTCGCTGGTGGCTCCGGACAGATACAGGT
TGCCGTAGCCCCCCAGGCCCTGGTTGACCAGCAGGGTGAACTGGTTGCGTTGCTTGTAGCTGCTGGAGTTCCAGGAGTCG
TTGTGCTCATCGGCGTGGCGCGCCGAAAGCGCGTCGCCGAGGTCGCGATAGCCCTCGGTGGAATAGCGGTAGCCAGCCAG
GGTGAGAGTGGTCTGGGTCGGCTGGAAGGTCCGGCTGTAGTTCAGACCGACACGCCAGCCCTGCTTGCGCTGGCCGTTCT
CCACCGTGGCATGGGAAAAGATGCTGTTGAAGCCGAAGGCTCCGTAGGGCGTGGCGAGCACGCCTCCGCCGAGCAGCGCC
AGATAGTCCTCGGCCAGGCGCGAGCCGAGGTTGGCGGTTAGCGAGTTGGTCAGGCCGCGCTGATAGGTGAAGTCGCCGAA
CAGGTCGTTGCCGTCGCCATACTGGCGCGCTTGGCCGAGGGTGGCGCTGTAGCGCGAGATGCCCGGGCGCATGGATTCGG
GAACCGCGGAAAAGGGCACGGAAAAGCGCGAGACGCTGCCGTCGGCCTCAATCACTTGGACATCCAGATCGCCGTCGTAG
GCGGTGCCATAGAGGTCGTTGATGACGAAGGGACCGGGAGCGACGGCGGATTCGTGGACCTTCTTGCCGTTCTGGCTGAT
GACCACCCGTGCGTTGCTGTTCGCTGTACCGCGTACCTGTGGAGCATAGCGGCGTTGCGAGTCTGCCAGCATGCGGTCGT
CGCTGGCCAGGCGCACGCCGCGATAACCCAGGCTGCCGAACAGATTGCCCTCGGTGAAGCTCTCGCCCAGGGTCAGTTCG
CTGCGCAGGCCTGGCACCGCGCGCTGGGCATAGGTGCGGATGCTGTTCCAGTCGCTGCGGGTGTTTCCCGCATAGCTGGA
GTAGCTGTAGTTGGACTGGTGGCGCAGGCGCCACAGGCCGAAATTGATGCCCCCGCTCAGCCCCAGGTAGCCATAGTCCG
AGTCGCCGTTGCCGCCTACTCCGTCGAAGCTCGAGCGATAGAAGCTGGCGTCGTAGTTGACGAAGCCCATGCTACTGCCA
GCGTCCCACTCGTCGGGACCCACGTAGCCGCGTGGCTTGATATCCAGCAGGGCCTGGGGCACCGAGAGATCGAGGCGCAG
AAGGGCGGTGTCGAGATTGAAGGTCGAGCCTGGCAGGCGTTCTTCCAGCCCCAGGCAATCTCCAGGCGCTTGCACGCCGG
CCTCAGAGGCGGGGCGGACGCCCAGCGTCCGGCGCAAGAAGCGTTCGCCGAAGCAGGGTTTGACGCCGCCGGCCCGGGGG
CGGAACTCGATCCTGGTGCGACTGGCGTAGCTGCCGTTGAGATAGACATCGACATGGTAGGTTCCGGGGTCCACCTGGTT
GGCCCGGTTGAAGCGTTCTAGGGTCCCGCCGGCGAGGCCCGAGCCCATCAACAGGCTGTCGTCGAAACGATACTGCGCCT
CTCCATTGGCCAGGCCCGCGAAGGCCAGCGCTAGGCCGCCCAGAGCCAGCGGCAGTTTGCCGGTGCGCCGTGCCTCAACG
TGGCAGAACATGCTGGACCTCCATTCGCGCGCCGTAGTCATTGATCAAGAGGGCGTTCAACTGGACCGGTCCTGGGGCGA
GCGGCGAAGGGCGTTCCGCCGCCCACTCGGCCTGGCCTTTGGGCGGGACCATGCTCGACAGCAGCCGCCAGCGACGTTCG
CCGATGCTCAGCATCGCGCCGGCGAAGGAGGCGTAGTAGCCGCTGGGGTTGTCCACCCGCACACGCCAGTTGCCGCTCGC
CTGCACCAGGCTGAAGTGCAACTGCTCGACTAGCTGCTCCGAGCCGCCAAGCAGCGCGGCGGGGCGATAGAACAGCTTCA
GTCGACTGCGCTGCATCAGCAGCATCTGGTTGCGCTCGGCCTTGGCCAGGTCGCGGGGCGGGATCTGCAGCACATTGAGA
TGGAACAACGATTCGCGATCCTCGGGCAGCGGGAGCCCGGTATACAGGAGGCGCAGGGTCTGGCCGCTGCCGGGGGCTAT
GCGCGTCACCGCTGGGCTGACCAGAAAGGGCGCGTTTGCAGTCTCCGGGGTGGAGGAGGGGTCGTCGTTGTCGATCCAGG
CTTGGATCACGTTGGGGAATGCGTCCTGATTGATCATCTGCACGGTCTTTTCCCGGGCGTCGCCGGGATAAATCACGCGA
GTACCGATGATCACCACGCTGCTTTGAGCCGGGATGGCGGGTATCAGCAGGGCGAGCGCCAGCGTAGCCGCACGGAGCGG
AGGCAGTGGAAAAGTATTCATCTGTTCACCTTGGCGCCGTGGCGCTCAGGAGGGGCCTCGGGTACTGCGGCGGAACCTGG
GAGGGCGCCGCTGGGCGGCGCCTCCGCTATTACTCGTAACGCAGGGTGTAGGTCACCGAGCCGAGCACCGGTCCCGGAGT
GACGGTGGTCGCTTCGGAGATGTAGCGGACGGCATAGTCGTAGCTGGTGCTGGTCTCGCCTCCTGCGAGCACGATGTCAC

FIGURE 2E

CGGCAGCCACGGCCGACCCGCCGGCCAGATTGACCTCGCTGCCGCCCCGGTGCCGTCAGTAGCTGCAGCGCTACGCCGGTG
GCGCCGGCGGACGCGGTATTGGTCAGATTGCCGGCGCTGGTCGGATTGACCGCCTGGAACATCACGCTGAAGTGCTCCTC
GGCATCGGCCGCCGGAGCGGCGCAACCGGTCAGTTGCAGGGTGAAGCTGGTTTCCCCGGCGGGTTTGCCGACTGCGCCGT
CAAGAGCGCTTACCGGTACGGTGTCGAGTATCACGGTCGGGTCGGTTCGTCCGTCGACGACGGCGGAACAGGTCTGGTCG
GTCACTTCTCCGTGGAAGGTGATGGTATTGGCAGCCGCCGCGATACCCGGCAATGCGAGAAATGCGGCGATAGTCAGGGC
AGTACTTTGAGGTTTCATTTCAGGGTTCTCCTTGTTGAGTTTTTTCCTCCGAGCAATGGTCTGTTAAGAGCTATCCGGGC
GTATTTAGGACGTTTCTTAAAGGACCGAAAGTGGGTTTCTCGCTCTCGGCAAATGTTTTTACTGCAATGGACTACCCCTG
CTCCAATAGACATTTCCGGGCCTCATAGCGAGGCGCTTTCAAAGGCCTCGGGACTGACGCCACCCAGATGACTGTGCCGC
CGGGTTCGGGTGTAGAAGACCTCGATGTAGTCAAAAACATCCGCCCGGGCCATGTCTCGGGTTTTGTAGATGCGTTTGCG
GATACGCTCTTTCTTCAAACTACTGAAGAAGGATTCCGCCACGGCATATCCCAACAATTGCCGCGCCGGCTCATGCTGGG
CTCCAGGTTGTTGGCAGCACAGAAGCGCTTCCAGTCGTCACTACCGTGCGGACTCCTATAGGTACTTTTGAAGGTGGAT
GCTGTGATGCACACCGCTACGCCGAGCAGTGCGGCGCAGGGCAACAGGATGAGCAGGGGATTGATGTTGATGGGAATTGC
CAGGTAGAGCGTCCAAGGGACGACCGCTAGCGGAATGATGCTGCCGCGCGCCTTGTGGTAGAGGTAGCTGGACTCCCGGC
CGGCGCCGAACTTGCGCAGGTCCCGGCGCACCAGGCCGTCCACCAGGCCGGTGAACGCGGCCATCACGAATAACGGGATC
GTCATGACCAGGATCACCAGGCGCACGACGAATGTCAGCACCGTGTACAGCGCGGCCAGGCCGTAGTTCTGCAGCGTGGA
GACACCCTGGGCGGTGAGATAGCGAACGTCCAGCGGGCTCCGTGGCCCGGCCTGCGCGATGGTAGTCATGTTGGTCATCC
AGTCGACCATCCCGGTCTTCACGAACAACCAGTCATAGGCCAACTGGGCCAGCCAGGTGGCGGTTCGACCTGGCTCCTGC
ACGACGACGCTGTGTAGCAGCCCCTGCGACAGCCAACTGAGTTCGTACTCAAACATGGCCTGGGCATGCTTCCAGCCCGC
GTCAGGCCAGAAGAAATACAGGCAAACCCACTCCACGATGATTGCGCCGATCAGCGACCCGAACATCACGCCGAGAAATT
GGAAGGGCGTACTGATGATGGTGCCGAGAAGGGTCTTCTGGCTCTCTTGTTGCTGCTCTGCTCGTTGAGTGACTTCAGCC
ATCAGGCGGCATCCTCACCGGTGCCGAGGTGACGGAAATCGTCGACCAGGTCCTGGGGCAGACCATCCTGCAGGGCGGAG
TATCCCGGCGCCTCCCACCACTCGCTGTTGGCGGCCTGGCCCTTGCGCATACCGGCAGCCAGCTCCTGCAGGCTTTTCGG
CATCACCTCGTCGGGGGCGACCGCCGGCAGCGGCATTCGGATCTTCCAGAGATTGCCGCCCTCGAGTAGCGCGAACGCTT
GTCCTTTGGGCAGACCAATGATGTGGGCCGGCTCGAGCATCGGCACGCTGGTCATCTGCACCTGGTCGTGCGAGCTGGAG
GTGAAGGCTACCTTCTTGTTGTTGTTGATCGCGTCGTTGGCGCCGCTCGCCGGCGTGCTGGTGTAGATCTGGACCTTGGG
GAGCTGATTGGTAAGGAGTTCGGCCGTGGCGGTCTCGCGCACCCGCAGCATGAACAGGTTGTTGAAGTTGCCGATGATCT
GACCGGCCTTCGCGCGGGAGCCGATCTTGGCCTCGATGTCGCTCATGGTCTGGGTGTAGGCCGTCACCTGCACGCCGGCG
CCGCCCGCTTTGTTGACCATGGGGATGAACTCGTCGCCAATCAGCTCGTTGAACTCGTCGGCATGCAGGTTGATGCGGAC
CTTGCCGCTGGCGAGCGAGCCGGGCAGGCCGTCATCGACACCATGCTTGTAGATGTGACCCGCTACCGAGACCAGGTCGC
TGAACATGGAGTTGCCCACCGCGGCGGCGACCTCGGTATCCGATAGTGCGTCGAGGCCGACGTAGACCACGGCGCGTTTG
CGGATGACCTGCATCCAGTCGAAGATCGGCCGCGGATCGTTGAGGTCCGCGTAGTTGGGCGAAAGAAGCTCCGAGATCCG
CCCGGTAGTGAGTTTCTCCAGTAGCGGCAGCAGCGAGGCCACGATCTTGTCGAAGTAGGTCTTGTCGTAGCGCACGGCGC
TCTTCAAGCCTTCCATGACCGGGTCGGCGATGCGTTTCTGTGTCAGGTACTGGTCGATGGCTACGACCCGCAGGGGCCGT
CCTTTCATGTTGAACGGGATGTTCTTGTCGTTGAGCTTGCCCTCGATCTGGATGATGGTGTCCCAGGCCCTGGGATCGTG
CTCGCTGATGTATTTCTGCGCATATTCGACGAACAACGCATCGATGTTCACGACGTGCCGGAGGATCTGCTGGTAGTCAG
GCCGGATACCCAGCGCGTGCAGCGCGCGGGCGATGATGTTGACGAACCGCCAGGCGAACTCGCGGAACGCCGCGCTGTTG
CCCTCGCCGGAGAGCTGGCCGGCGACGCGGGTGGCGACCTCGGAGATCCGACCGAACCGGCCGACGGCGTTGTAGCGTGC
CGACAGGTCAGGATGACCGAGGTGGAACACGTAGAACTCGTCCAGGCGGCCGGCACGTTCGCATTCCACGTACATACGCT
TCAGCAGGTCAGCGTCGCCTTTCGGGTCGAAGACGATCACCACCTCGTAGTCCGGCTGCTCCTCTGCGCGCCGGCGCCGA
TAGCCGTGGTGAACCGTCTGGGTCCGCCGGCCCATCTTCACCCGCCGGCGTCGTACCCGGCAGTGAGTGCGGCGAATATC
CTGGGTGATGAACAGCTCCGCGAGGCGCGTCTTACCCACCCGCGTGGTGCCGAGTACCAGTGTGTGGCCGACGCGCTCGC
CCAGTTGCAGGCCTACGTCCTGTTCGCGCGGCTCGATGCCATGGAGCCGAGGCAAACCGCCCACGGGCGGCAGCGGCCGT
GCGGGGTTCCAGGCCACGTCCCAGGCAGTGGCTTTGGCGACCAGCTTCAGGGGGAAGGGGGCGAACTCGAGCTGCTTCTC
CAACCGGCGCGCGCGCTCGTAGAGGGGCGAGGGCTCGACGTAAGAGGCGAACTGGGGCAGGTAGGTATCTGCCAGGCGCT
GCGTATGCTTCTGCGTCCAGCGAAATCCTTTACCGATGAACAGGTGTTCGTTGCTGACCGGCATCTCGGCGCTGGTCATC
GTGTACTTCGGCAGTCGGCGAATGTTCCGCCGGTAGCGGAGCACCACGCCGGCCTGCTTCAGCCGCACGATACCCAGCCA
GGCGAAGCACAGCGCGGCCACGATGCCGAACAGCGGAGTGAGGGAGAACGTCCACGGCGCGACGATGCAGAGCGCGGCTG
CGGTGAAGCACACGGTGGTGGTGTAGAGCTCCACGGCAGGCCGCAAGAGCGCTTCCAACGGGTACTGGCCAGCCATGGCT
TACACCTTCAGGCTTGCCGGCGCCTGCAGGCGGCCCTTGGCATGAAGTTCGAGCACCTGTTTCCTGGCTTGCTCTACCTC
ATGATCAGGCAGCATCTGCACCAGCAGCTCGCAGAGGGTCGACTCGTCCTCCAAGGCCTTGGACGTCACCTCACCGCCGT
AGCGTGAGATCAGTCCGAGCAACTCCTGGTACATGAGTTGGCGTGCCTGAGCGTCGCTCGGCTGCTTGATGACTAGGGTG
ATCAGATGAGTAGTCATGAGGACTCCTACTGCTGGATGGCGGTGGATGTGATGAGGGCGGGGTAATGCTGCAGCCCTAGC
CGGTCGACCAGGTCGTCCGCCGGCGCCGGCAATATCTGAAGTCCTTTCCCCCAGGCCCGGATTTCCGTCAGGCGCGCTTC
GCTGGCCACGTTCACTGCCAGGCCCACGGCTTGGAGGTCTCGAAGCTCGTCGCCTCGCTCTTTCAGCCAGGTTCGAGACA
GCGTGTCGTCGCCGACCAGGAACAGCAGTTGCAGGCCTGGAGCGTTGATGGCGCGCCCCTGGACCCGTCCTGGCGATAGG
CGGGCGGAGCGAACTGGAAACGCACCTGAGCCACGGACGCCCAGGTCGCCAGGTCCTGGTGCGGAAGCGGATGGCTGCGG
ATCCAGGCCCTGGTAGTAGGGGAGCGCCGAGGCGCCGCCGAGGTCTTCGACAACGATCAGCCCGGCGGCCCAGGCTTGAG
TCGCGAACGCCGTGCAAAATGCCAGCGTGGCAAGGAACTTCTTCATGGATTGGTCACCAGTAGGTTGACGCCGAGAACTT
GGCTAAGGTGGCGGGAAACCGTCCTACGGTATTTGGCGGCAGGCGCGCCGCCGGCGGGGCGGTGGTAGCGGCCGGCAACT
TGGATCCAGTCCACCGGTCTTCCCGGGGCGGATGCATACAGGGCGTCCCGCTGCTCGATCAGGATGTCGGAGGTGGCGTC
CAGGTTCTTGTACGGATCCAGGGACTCGCAGGGGCTGGAGAAACGATGTCCGTTCCAGCCGATGTTGACCTGGCCGAGGC
CGGAATCGACGCTCTTGGCCCCGTAGAGGTTGATCGCCGCGAGTAGCGCTGTGCAGGCGGCGGTGCGGGTCGCGTAGTAG
TAAGATTTCCCGGCGACGTTCAATGTCCAGGGCCAGGGAACGTATTCGCCGCGCAGCAGGACCTTGCTCTCGGTCAGCGC
GACCGAGTAGAGCACCTCGGCCGGAACCCCTTTGGGCAGGGCGATTTGCTTGTACGCCGGCGGTGGATCCACGGCGGCCT
GAGCGATGTTGACCAGGACCAGCAGGGTGGCCAACTGGAGGGCTCGAACTACAGACGTTGCCATCGGCCGTTCACCTCCT
GTACCAGGGCGGGAAGGGCTCCGCCCAGTCCTAGGGCCATCCAGCGGCCCTCGTCATGATTGAGCGTGATCTGCTTGCTG
CGAACCTTCTTCGGGTCGATGCCGGCGAGGATTGCCCAGCGCCGCACTCGCTCTGCGTCGTTCTGGCTACCGACGAAGTA
GAGGTCGAACTCCTTTTCTGCATGTTGCAGGTCGCGGACCCGCTGGATGCAGGCGGTGCAGTTGTCCTGGACGAACAGGG
CCAGGCGCCCGCTGCTCTGCAACGCTGGGCTCATGTTCGGCGTGCCCGACGGGTTGGCTGTGCTGCTTTCGGTGAGGCGG
ATCACCCCCTCGCCTGGATAGGCGCGGGCGAAGGCTTCGTCGTATGCGCGCTGGTAGGCGAGTTCCTTCTCGACCCGGCG
CCGTTCGGCCTGGACCTGTAGATCGGCATACCGCCGCCGTTCCTCTGCCGATCGCGCCTCGATGCCCAGCGCGGTCAGCG

GATCAATACCAGGCGAGTAAGCGCCCCGGGGGCCTTGCATGAGCGTCTGGTAGCGGGTCCATTCCTGTTGGTTCAGTCCC
CACTCCTGGGCCAGGTCGCTCTGAGATGTCTGTTGTTTCAGCGTGCCGAGTTGGCTACCGGACAGTTGGCTGCTGGCCTC
TCGGCTTGCCGCCGCCTCCTGGCCGGCTGCGGGATAGCTCAGCAGCAGCATCAGGCCGGACAGGAGCGATACGGTTCTGA
TCATGGGCGGGGTCCTATGGTTGGATGCGAACGCTGCGCGAGGTGCCGGCGACGTCGAAGTGCGCGGTACCATCGTCAAG
GTCGGTCAGTCGCCAGGTCGTGCCGGCGACGGCATCTCCCCGGCGAATGAGGTAGATCTGGCTGAGCTGGGTGGATCCCG
GAGGTGCAACCGACAGAAACCGTTCTCCCCCGCGATACTCCACGCCAAGGATCGAGAAGGGCGGGGGCTCCATCGGTTTG
GGTTTCGGCTTGGGCTTGGCGGGTATGGGGCGTTTTGGCGCTGGCACGATCAGCGGCGGCGCCTTGGCCTGCTTTCTTG
AAGCGTGCGGACAGAACCGTCCAGTGTCTCCACGGTTGCCTTGAGCACCAAGAGGTCGCCGGCCGAGGCGGTGGTCTGAG
CCAGGTTCTCGACGGCATCGGAGGCCTGCTTGGCGAACGCCTGCGCAGCGTCAATTCGGTTGGACAGCGCCTGCTGGCCT
GAACGGAAGTCCTCGTTGCTGACCAGGTGCTGTCCATCCACGGCGTCGAGGCGCTCGTCGACTCGACTCAAGCGAGCCAG
GATCGCCTCCAGCGAGGCCTTTTCCGCGGCACTGTCCACGCCTGATCGGAGCTGAACGAGTTGGTACTGCTGGTAGCTCA
GCAGGCCGGCCATTGCCGTCAAACAGAGGCCAAGAATCATTGATGCAGGGGATGGGCGTTTCATGAGGGTCCTCGCGCGA
AGCTGGGGACAACCTCTCAGACGGACAACGTTGATGCAGTAGGAAATACTCGGCGGAAGGGGGGAGGATTTTTTGAGGGG
AGAAATTCTTGCGAATCCTCCGGCCTAATGCTCTGTGCGAAGCTGACGAAGCTCGGGGGCTCTTCGTCGCGCATCAGGGC
TGAACGCTGGATAGCACAGCAGGCTGACCTGCACCGAAAAAACCAAGTGATTGGCATCGCCGCGTCGCTGGCGGATCGAC
GATGGGGCGGACTCGTTGAGGTGGAGTAGTCAGTGAGGTCATTCGGCTTACCGGCAGTGCAAATAGGATATCCCGCTCCG
TGACCGAGAGGATCAAGCGGCGCGGCATGTCTCAGATGTCTGTGCGGGTGTGGTGGCAGAGAGTTTCCCGGCTGATATTC
GGAGGTGGGGAACAATAGAACCAAAAACCAACGTAAGCAGAAATGAGTGCTGCACCCAACAACGCCGCCACAATGCCACT
ATCGATTTGACGATACCGGTATACGTAAAAGCATGGAGAATCAGTATGTTACTCACCCCAGACGAAGGTTAGGTTGACCA
TTGCGCTGCCTTCTGAAGTTATGGGGCATAGATTATGGCCGAAGCTATCAGAAAGGATGCAATGATGACAAAACTCTACT
TTGATCTTCTGAACTCGCCTGCCGAGGCTCATTCGTCGATACAAAAGTCTTTATCTGTGCAGGCAATCTCCACAACTGTC
CCAATACTGGAGTTTCCTTCGGAAACCGTATACGCCTATGCATCGTACATAAATGCATTAAGTATCGGTCAACGCATAGA
TCCTGCATTCACCCAGAGCTTAACGAGTGCCATATCCAACCTGGCAGGTCGCCCGATTGCAGTAAGCGACATTTACCAAA
AAATTCATGAAACCACACTGAGAACACCTGTTGAGATGGGCGTTCGTCCTAATAGCATCACCTTTGAGGAGTATCAGGCC
ACCATAAATCAGCAAGCCATCAACATGGTTCAAGATATGCAGGATGGAGACAAAGGTGAGAAGGTGGAGGCCCTCCAGGC
CAATATGCAGTTCCTGTATGGACAGGAGATAAATACTGATTTCATCGCTCGTAATGAACTCGCTGCTGGGCAGAGAGCGA
AAACCGTCGCAATAGTTCAGGGGCATATCACCATCGGGTACGGCTTCGATACCTTCGTGCATGAAGCGTCCGAGCTAAAC
TCTTTGAATCTTGTTGGTTCTACGCGACAGAAGGTATTACCTGCATTGCAGCTATCAACGTCCGACCCAGGCTTCTGGAG
CGTCTATGCCTTGCTGGGACAAAGTCTCACGGATGACGATGGGCTATTACTCTTTAGTGCCAAAGCGCGAGCTGTTGTTC
AACGCATAGCAAGCAACCAGTTTGCAGGTAAGTGGAATGGGCTACCCCCAGCTATCAAAACGGTTGCGCTTGATCTATAT
TATCAATATGGGCAGACTGGTAATTTTCCAAAATTTCAACAAGCTATAAATAGCCATGATTGGCCGGCAGTCATCCATGA
ACTTAGAAACTGGAATGGTGTACCGAATGATCCTCTCCAGTTCATTACAAAACGATTGGAAGAGCGAGCCAAGTATCTGG
CAATATCCTTCAACTATGAGCAATGAGGATATGGATATGAACAACACAGTGAGCGAAACGCAACAGATCAATATTTACCA
AAATCCGGGGCAGTCTATTTCCGGTCTCTACAAGGGGCTGGCTAACCAGTGCTCTCCTGGCCAGCCATTTCCAGAGGTAC
AGCTTGTGGAGGCTTGGGATATCCCTCTCGTACTCCATCCGGAGTTTGTGCCTAACGGAGATGTCTCGAAAATCGATAAG
GAGTACGGAACGATCCTTGCTGCTGAGTCAGCTCAGGTTATCCTGCTTCAACTCCAAATGGCTCAAGACAAGGCTAAGGC
GTGCGGGGAGGTTACAGCCTTGATCAGTTCTGTCTCCTCCAATCTCAATACCATTAAGAGTCGTCATGGTGCTAATTATC
TAAACCTGCTGAAACAATCACCGAACCGATACCCGACTAGCGTCGGAGTTGAGATCATGTCAGGTGGCAGTCCGAACCAG
GATTCTGGAATCGAGGTCTCTTACGGTGCCAGTCTCGGCCGTCTAACTCAATCACAACTTCAGGCGATGAATCTGCCTGC
CAGTCTCAAACAGTTGCTCACTCAGGGAATCGGTGTGAAGCTTTCTCAGCCTGAATATTGGCCTGCTTACAACAACATAG
CCACTGGTATTCGTTATACAACCGGAGTGGCGATAACGTTGGCCTATTGGGCCACGGTTTAGAGATGGTTGAACATTAGG
GAGAAAAGGATAGGTGCCCAGATCAATCTGCTCGTTCATTTTTTTGCTCATATTCGCGCCCACAATGACCCAAGCTGCGA
AAATACCAGCAAATGAGTACTCATTGGGGGATGGAAGAGGCTACATCAATATCTGGCCGGAAAAGGATGAGGCTCAGGCA
TTTCTTATCCATAATGATGGGCCTAATGGGGCTACATGCAGCCTTAAAGGCACTCTTAGAGATAATAAAGGAGTGGTGCA
TTCGCCGTATTCCTCTGCTTCATGTTTGCTAAGTATCACCCAGACAGGGCTGCTGTCAGTAAGCGTCAAACGTGAGGAAA
ATTCGCCAAGCTGCTCTGCATGGTGCGGTCCTAGAGTTTGGTTTGAAGGAGCCTATAGCGTCCCGCCCAAGGGCTGCTAC
TATATGCAAATAAGGAAAAAAACTCGACAAATGTTGGGTATGATTGAGAAAAAAGAGCTTGATGCCGCTCGCGCCTTATC
AAATAAGCTTTTGTCAGACTGCGCAACCGAGCTAGCCTATCCTGCCAAGATATACTTGACGAACACACTTGCCATGATCA
GTGCTGAAAAGGGAGAGAATGCTCGCTGTTTGGAGTATGCCCATCGGGTGCAAAAGCAAATTCCTGTAAGAGATGACGGC
CAACCGGCTGAAGACTTGCTCCCGGCGGAGCACGCTTTCGCTATGGAACAACGCGCCAAGGCTGATGCTCTGTCTGAGCG
ATGCAGCGACGAGAAATAATGACGCGTCGTAGTAGGCGCCTAGTTCACTAGGAGCGATCTTTTCCGACTCTCCCGCGGCA
TGAGTCACTCGTTCGTGCGGAAGTGCTGGTAGAGCGTTTGCCGACTGATGTTGAATTCGCGGGCGAGCTGAGCCTTGGGC
TCGCCGGCCGCTGCCCGCAGCCCCAGGGTAGCACCTGCTTGTCGGACAAGGCCTCTTTGCGGCCCCGGTACGCGCAGAGC
TTGATATCCTCGCGCTACCGCGCTGGTGCTGCTTGTATGCTGCTCAGCAAGCCAGCCGCCGGTCTTTTCAGGGTAAGCGT
ACGGCCAATACACCTTTACTTAGGTTGAGCCGTATCAGCAAAAGAATCGATTCATATGTGTGAGGTACCGCCTCATGCAC
ATTGGCTTTGGTACCGACATCTACTCGGAGACCCTCGTGCCTTGGTGAAACACCATGTGCCAACCATCTTCGTAGGTTTT
CCAAACAGAGCTACGCATTGATAACGACTCATCGCCATTAGCATTTTGATGACGGCATCGGTACGTTGCTAACGCGACAC
CATCTGCAAGCACACGCAGTTTGAAATCCTCGATTGTCCTTTTGATCCAAGTCTGGGATTTTAGTCCCACGATCACCTCC
GCTTTCGTCCAGCTTTTGCCGATAGCTCCAAATTCAACGAAGTCTTCGGTAAGCAGCTGTTTAAGCAGTCGATCATTCTT
TCTCGATTCCGGCTCTAGCAACAGCTGTTCAAGCGCTAAAATCTCCAGGCGAATATCCATTTCCCTCACCTACCAAGTAT
GATTTATATATGTTTCACAGGGGACAATAGATCACATAAGCTAGATCTGTACCCACTGATGCGGCAGCAACTGGCTGATG
TCTTTCGACCGTAACGTCGGCAGGCGAGTTAGCACGTCCTTCAGGTAGGCATACGGATCATGCCCGTTCATGCGAGCGGA
CTGGATCAGGCTCATGATAGCTGCTGCTCGTTTGCCACTGCGCAGCGATCCGGCAAATAACCAGTTCGAGCGTCCGAGCG
CCCACGGCCGTATCTGGTTCTCCACCTGATTGTTGTCGATGGGCACAGCCCCATCGTCCAGGTAGCGCGTCAGCGCTCCC
CAGCGTTTCAGGCTGTAGTCGAGGGCTTTAGCTGTGGCCGAGCCGTTGGGCACCAGGTCGCGCTGGGCCAACATCCAGTC
ATGCAGTGTTTTGCTGATCGGTACCGCCATTTCCTGACGTATTCGCCAACGGTCTTCGTTGCTCATGTCCCGAGCCTGGC
GTTCAACCTCGTACAAACCGCCAATTGAGTGCAGCGCCTGTTCGGCCAGTTGGCTTTGTTAGCGACATGCAGGTCGAAG
AACTTGCGGCGAGCATGAGCCATGCAGCCGATTTCAGTGATGCCTTGTTCAAAACCGGCCTTGTATCCAGCGAAGTCGTC
GCAGACCAGCTTGCCATTCCAGTCGCCTAGGAAGTTGCGTGCATGTTCTCCGGCACGGCTTGGGCTGAAGTCGTAAACCA
CCGCTTTGAGCGCCGAAAACGGCGTCGTGCTGTAGGCCCAGACATAGACCCGGTGGGTTTTCTTCTCGCCTGGTGCAAGC
ATTTGCACCGGTGTTTCATCGGCGTGGATCACGTCCTGGTTCAGCACGGCTTCACGCAGTGCATCGACCAGTGGCTGAAG

FIGURE 2G

CCGCACGCCAGTTTGTCCGACCCACTGCGCCAGGGTCGAGCGGGCAATTGGCAGCCCGGCGCGGCCAAAGATTTTTTCCT
GTCTGTACAGCGGCAAGTGATCGGCAAACTTGGCCACCATCACGTGGGCCAACAAACCTGCGGTCGGGATGCCTTTATCA
ATAACCTGGGCTGGCACCGGCGCCTGGATCAGGGTTTCGCACTGACGGCAGGCCCATTTGCCCCTCACATGTTGCTCGAC
GGTAAACACGCCCGGCGTGTAATCCAGCTTCTCGCTGACGTCTTCGCCGATGCGTTGAAGTTGGCAGCCGCAGGCGCATT
GGGTATTTTCAGGTTCGTGGCGAATCACCGTGCGCGGGAACTGCGGCGGCAATGGCGAACGTTTCGGGGATTGCCGTGGC
TCGGCTTGTGGCGAAGCTGGAAGGAGTTGTTTCAGCTCGGCCTCGATAGCTTCAAGGTCGGTGTCGAGCAGGTCATCCAG
CAAGCTGCCTTGCGCCGAACTGATTTGCTCGCTGCGCTTGGCAAACTTGTGGCGTTTGAGCAGAGCGATTTCGAACTTGA
ACTGTTCGATGAGGGTTTCATTGTTGCGGATTTTCCTGCTCATCGCCTCGACCTGGGATTGCAACTGCAACGCCTGTGCC
GCCAAGGCACGAAGCTGTTCCGGGGTCATCTGGTCAAGGTTGGGCGAGGAAGTCATGCCGCCGATTGTGCCAGAGCAGGC
GCGATGCGACGACAAGTAGACCAGCCAAATGGCCGGCGTTTACAGCATGCTGATCACACCGCCTGTGCCGACGCGCTGCC
ACGGCAGGCCCAGCACCAAGGCCTGGAGTTGTTCGCTGTCGAGTTCGACCTCGCAGCCATGGCGAATGCCGGGCCAGTGG
AACTTGCCTTGGTTCAGTCGACGCGCGGCAAGCCAGATGCCCACGCCATCGTGCACCAGCACTTTCATTCGGTTAGCCCG
GCGATTGGCGAACAGATAAGCGCAGTGCGGCTTCGCCGCACCGAACACCGCAATTACCCGGGCTAATGCCGTCTCGGTGC
CGGCGCGCATGTCCATCGGTTCGGTGGCTAGCCAGATCGCATCGATGCGGATCATCGAGCGAAAGCCCGGATGAATTGCG
CGCAGCCCTCGGGATCTGAGGCTGGCTATTTCACCGTGATCATTTGCCCGGCCATGGGCAGTTCAATGAGCACTGACGTT
TCGGCTGGCCGTTTAGGGGTGGCCTTCAGCGGGACAAAGGCTGGTAACGAGGCTGGTACGGGCTGGTCTCGATAAAGCGT
CAGCCATTTGCGGATGACATTGGCGTTGATGCCGTGGCTGATGGCGACACTGGACACCGTTGCCCCAGGTTGCAGGCATT
CCTGAACGACCTGGGCTTTGAACGGTTTCGGGTAAGAGCTTCGTTGGCGCATGGAAATCCTGGCGATAAGGGTGATCGCG
TCCGCTTAAAAATACGCGGACACCATCGCCCTTAATGCTGGAGTTCGGAAGGTGAGTTCGCCGGACGCTTACTTTCAGG
AGACTACGAGTCATGTGCCGATGCAGTAATTGACTTGCCAATGAAGACAGTCAAGATTGACGCTTTAGTAATAACGGAAA
TTATCATGACATAGGTTCTAAAAGAGGCTCCATTGGAGCCTCGTATTTTTGTATGGTCGGCGTAGCACTTACTTTAGTGA
TACATGCTTACCCATGAACTGTATTCCGTTAGATAAGGAGCTGACTTTAAGTGAGTTCTTGTAGCCATAGAACTCCCCAA
GACCATCCCCGTCGGGAATTTCAGTGATTCCATTTGCAATGACCGTATGGCCGCCTAAACCTAGAAGCTTGGGAGCATAG
GGAACTGCTCCATCTGAAACATAAAGGGTGGGTGAGCGATGGTCTCTATCTTGCTTCCAAGAAAAGTACGAGTGTATAAC
TCCTGCTAGTACCTCGCTAGCCGGATCATTCCTGGAGAAGAGTACGACATACGGCCGCCATCGACTCTCATGCTCACATG
AAATGCCTGTGTGGAACTGTCGATTCTCTAATTTTTTTATGAGATAAAGTAAGGTTTCCTCTGAGCTTTTTTCGAGAGGA
GGTCGGCGTATATCAACTTTAGATATTTCGCCAGAGGAGGATGGCTGCATCCATAAGGGAGGGTATGCGCCTGAAATGAG
TTGATAGATCGCTGACTTTATGGAGAGGAAGGGGACGCAAGGGAAGTTGGAGGAGCCAAATACTTGATTACGAGACTTTG
CATGCTCATATGCACTGATCTCTAAGTGAGTTGACCATCCTTTATTACGTATTTCCTCTTGGATGTCGTTTGGCAATAAA
CTGACAATTGACTCGATCCAGCAACGAGCCTCGGTACTAAGCGAAACGGGAAGGTTTCGGTCTCGCAAGTCAAGAAATCG
AAATTTATTTAGAGTTTCTTCGACTTCTAATTTTTCTATTGCGTCGAATATATGGGTTGTTTGGATTGAGGTCTCTTTTG
TTTCGATGAGGATTAGTCCGGGTAGTTTCTTAAAGCTCATTTGCTCAAGAAAAAGGACTAAGGCTGGAAGTATAGAGTGC
TTTGGGTTTCCAGTAAAAACGACGGGGCCATTCTCGGTATATTCAAGGTCTAAACAATATAGCGCTTCCAGTACCCCCTG
GTGTGTATCGTTGCGGCCACGGATCACCCAGAATGGAGCGCTTCTATTCATGCCAATGTGCTGCTCCAGTTTATCGACAT
CAGCGAGTGAGCGGAAGCCGCGAGAAACTGCTACCGCATTTTGGCAAACCTTGAGTGGATATTCTGTTCCAAACACAAGA
AAGGAGTTCTTCTGCAAGCGTTTGGCATCGCGTTTCAACGCCGAGATGTTTCGTGAGGCGGACGCAGTGAGACTACGAAT
AAGAGATAAAGCCATGTTTTCCCCTTGTGCCGTTCGTGGCACTTACGGAGGTAAGGACTGGCTCGCTGTGGCACTTAGTG
CCAATTTTGCGAGGTTCATATAGTGACAGTGCGTCCCTCTCGGGTTTCATGCGCCGCACCAGCGCGGCCAAGACTGTACC
CCATTTATTAGATGACGCACAAATCTTGATTTTTTTATCCTGCACTTCAGTGCCTATGTGGTGAGCGTAGAGTTATTTTC
AGCTGTCTCTGGTGCTGTTGCTAGGAGATGCAGGTAGTTCAGACAAGATGTGGAGAAGGACTCCCAGAAAAAAGGCCACC
CGAAGGTGGCCTATGAGTTCATGCTGCGATGAGTTGTCGTGCCAACGCCATCTCCACAGAATCCCCATCCTGGTTCAACG
AGTCGAGACCTGACTCGGGAACGTCTCCCGACGTGCTCTGAGCCACAGCGATCTTCTTGGCCATCAGCTGTAAGCAGGTG
ATCTGCGAGCTGCCGGCGTAGCCGAAGAACACCACCCTCACCGGGTGCTTCTGCCCGATCCGCCACGACCGCCGCGCGGC
CTGCTGCAGGGTGTACACGTTGTAGCCCGTTTGCAGGAACGCGATGGTCGGGAAGTCGAGCAAGTCCAGCCCGGTCTTCA
CCAGCTCCGGGTTGGTGATCAGCACATCGATGCCGCGATCGACCTGGTCGAGGATCCAATCCTCGCGTCGAGCGGTATCG
ACCGAAGCACGTAGCACTGCCACCTTCAGCCCGGATTGCTCGAGCACTTTCTTCAGCCTGGACGTGGTGTCGCGCGTCCC
GCTGTAGACGGTGTATGCCAGAACCTTGCGGCCCTTCGCTTTCTCCTCGAAGCAGAGGTCCACCAGCACCTGCTCCTTGG
GTATCAACTGCTCGTCACCGAAGATCGCTGGCACGAAGGCCAGTGTGTCCCGGGTTCGCGGATGCTTGACGATCTCCGGT
CGGAAACAGCAGTCCGGCCAAGCCAGCAGCACGTTGAGGACCACGCCCAGGAGCGTGGTATCTCGTCGCGCCAGAGCCTG
GCGGAGCTCCGCTGTCAGCGTGGCCGCCAGGCGCTGATAGGCCGAAGCCTGCTCAGGCGCCATGGGCACGTCGATGAACT
CCTCCTGGTAGTCGGGCAGTACGTTGCCACCAATATCCTTGAGCTTCAGGAACACGGTGAACGGCAATACGAAGCGGTGG
ATGCCCTTTGGGCCGAAGCCGGGAGCCTTCACCGTGCGTACCGAGAGCTTCTTGCCCCGCGCTGTCTTGTGCGAATCACC
GTCGCGCTCGGTGTAGATATCCTTGAGCACACCGTGGTCGCGCATGAACGACATGGCTGCGGGAGCCATGCTGCCGCGCG
CGTTGGGCCGATAGCCGTCCTCGATCATGCGCTGGGTGAGGATGCGGAACAGGAGATAGAACAGATCGTCGGCGTAGCCG
CCCATGAGCGTTCCGGTCAGCAGCACGGTTTTCCGTGCCTTGGCTGCGAGAACGCCCATGGCCTGGCCCTGGGCCGAGCC
GCTGTTCTTGTACTCATGTCCCTCGTCCAGCACCAGCAGGTCGAAGTAGCCATCAGGTAGGTAGCGCTTGATGAACTCGG
TCGGTTGGTAGCCGCCTTCACCGAAGCCGAACTCGATGTTTGCCATCGATCGCTCCATGCGTTTGGCCTGCCGATCGCTG
AAGACGAAGTTGCCCTTGGCGTCCATCAGGTTGATGAACTCCGAGACGTTGTCCACCAACATCGTGGCCAGGAAGTCCTC
GCCGAAGTCGTTCAGCAGGCGCTCCGCCCTGACCGGGCCGATGGTTGGTATCCGGCACATCGACTTGAGAATCGTTGCGC
GCCGGTTGCCGCCGTCGGGCTTGCCTGGTCGGATCAGCGTCCAGAGCGCCCCACGGCAGGAGGAACAGGTACGTCGACGG
TCACCACGCTCGAACTCCTCCACCGTGACCAGGTTGCCTTCCAGGTCCTCGAGGACCTGTCCGCAATCCGGGCACGCAGC
GAGCAGTTGGCCGCCGGCGGCGCGTTTCTTCCAGCAGGCGAGCCGCCAGTGGAAACCCATCCGCATCCGCACGCGGCCGA
GGATGAAGAACTCCTGGCGCCCGTCGTAGGCGTCGCCCATCTGATCTCGCAGCTTGAGCAGCTTGAGTAGAGTATCTGGG
CCATTGAGTACCCAGACGCGGGCGGCTGGGATGGTCTCCAGGATCTCGCGGCGCCACTTGTAGACCAGGTGCGGCGGAGA
GACGACCAGGGTCCGGCGATAGCCGGCGGCGTGCATGACCGCTGCGACAGCGATGGCCATCATGGTTTTCCCGGTGCCCA
TCTCGGCGTTGATGATTCCGGCCTGCTCGTTACGGTCCAGCAGCAGGGCGGTGATGGCCTGGACGACCTCGGCCTGGGCC
GCGAAGGGCTTGCGCTTGAGTCGGTCCATCACCAACTGGCGGTGAGCGTTGACGGAGCCGGTATAGACCGGGGGGTTGGA
GCGATTGAGCGACTCCAGGAGCTCGTCGCCGAACTCGTCGATGAAGTCGGTCAGGTTGATGTTCAGGTGGGAGGCGGCGA
GGGCGGCCGGTTGGGTCAGAGCGTTCATGTGCATCTCCTGCAGGTGAAGAACGGAGCAGGGCGTGCACGGTCCCCTTTGG
GTATCGGCATACCCTGCAGGGTGGAGTGGAAGGAATCGCCGCGGGGGCGAGGTCAGTAATCGCTAGGGAGCAAAAGCGTG
GTTGAGCTGCGGTCTGCCTCAGTGATGATCCAGAGCCTGGATTCGTCGCCGGCGTCGATGTCGTAGGACGACAGCAGCCG

FIGURE 2H

ATCGCCGAACTTCAGGGCTCTCTGGTTGGTGTTCCAGTCTTCCTGGTCCAGCGTTCCCCAATCTCCCGCCAGATGGCGTT
TCAGCAAAGGCGCTGGGTTGAGTTGACCGGTTTCCACCAGGTGGCTCACGGCTGCGGTCATTACGACCCGGCCAGGGCTG
ATCAGCAGCGGTGCGGGTGCGGGGGCCGGCTCAGGTTGGGGCTCTTCAGCTTCTTCCGTGGTCGCGGCCGAGGAGCTGAT
GGTCAGCACGCGGCCCTGATTGACCGAGGAGGGTGTCATTTCCCATGCCCGGATGATCGGGATGAAACGGTCGGTGAGGA
TCCTCACCTCGGTGATGTTGCCGTCGTCGTCCTCGGTGAATTCGGTCTTGCGGACCTTGTCCTTGTAGGTGTCACCCTTC
ACGACCAGGATCCGGCCCGACTTCGATCGCACGACGCCAGATATCGCGCCGGCGGCCAGGGCCAGGGCCAGGTGCCAGCG
AGACAGCTCGCGGACTGGTGGGCGCGGCTGCAGCCCCGCTTGCGCGAAGTGCAGGTTGAAGTCAGGCCAGAGACCTCGCA
GCCGCTGGATTTCACCGGCGAACTGCTCCGGCTCCAGGGTTACTCGGTAGAAGTGCTCCAGCTCGCTGGTGGCCGGCAGA
ACCACGTAGGGTTCCCACGGCCAAGCCGCTGGAATTTCCTCGGCCTTTTCCTGGCCCGCTCCGATCGCCTGCAGGCGAGA
CCTCACCTGATTGGCGTCCGCCCGGGCCAGGTCCTGCCGACGGACCCGGATGCCGAAGATCACCACCTGCTTGAAGGTAG
GATCCGCGGCTGCGTAGATGCGCAGGCCGGTGAAGTGGTTGCTCAACCAGCCAGTCAGCTCATCGTCCAAGACGTAGTGA
GGAACAATCAGAACCATGACGCCGCCGTACTGCAGCAACGGCAGGCAGCGCTGGTAGAACGCTTTCTCCAGACGCCGGCG
GCCGCTGCCCTGGTACTGCGACGCACCGGAGTGGTCCGCCACCAGGTCGCCATAAGGCGGGTTGAGCCAGAGCAGTCCGA
ACGACTGCCTGCTGATCATGGTGTCGAAAAGGTCACTGTGCAGCACTCGGTCAAGCAATCCTCGGGCATGGTCGGCGCGC
TCGCGGTCGTACTCGACAGCGAGGGCTTGGACCTGATCGCGGCCGAGGGTGTGTGCTGCCTCAGCCAGGGCAACACCCTC
ACCGGCGCAGGGGTCACAGATCCTCATCCTTCCCGACGGGGCAAGAGTGAGGGCCTGCAGAGCGCGTTCGAGGGTGACCT
CATCGGTAGGGAAGTAGCCGTTGCGTGCAAAGTTGCGCGCCAAGCGCGGGAACATGAGGGCCATGGGGCCTCCTTGAGTC
TGAGAGGGATGGCTCAGGCTTGCGCCTGAGCATCGGTGGTAAGAATGCTTTCGCGGATTACCTCACCGAGGGTGGGCTCA
AGCACGTCGACCCGCAGGGCGAGTCGCCAGGCGCAGACGTTGCCGATCGTCCCGGGTAGGGCCGTCAACATCTGGTGCTG
GGTGAGAACCTCCATCACCGGCTCGCGCCAGTGCTGCAGGAGGGGGAGCGGACAGGTTTCCATCACCAGCGGCCAGAGCC
TCTGGTGTGGATCCTCATCGCGCTGCAGGAGTGCGAAGGCGAGGTGATTGCCTCGGTCGGGCGCCGAGGCGCGCCGATCG
AACAGCCAAAGATTGAGCAGGCTGCCGAACAACGTGCCGCGGAACTGACGGGTGGTGCGTTTCTCCAGGAGATCCTGGTT
TGGGAAGACAGGTAAGCGGCGGCCGTCCACGATGATGTGGAAATGGTCGATGCCATTTTCTTCCCGGCCCAGCGTCAGCC
TGGCCAGGAACTCTTGTGTCACGGTGTCGCGGCCCCAGGCCGAAAGAAAGACCAGGTTGCACTGCTCGTCGCACACGCAG
GCGTCGACGTACAGGTCTGGACACTCTTCGATCTGGTAGAGCGGGGTGGGACTGGGCATAGGAACCTCCTGGAAGGAGGA
GCCACGCCGCCCTCAAGGGGCGGTGAAGCCCCTCGGGGTGTAGTCCAACTGGTGCAGGGGAGCGTGGTGGCTACCGAACG
CTCTTGGTGTGAGGGTCGAAGCTGAGTCCATCGGCTTGGTGCAGCGGCCCTTGGCCGATCAGGAAGACCTGGCAGAGGTA
CTGATCGCGTAGCGCGACAGCCGCTTCCGCCTGCTCCAAGGTCAGCGACAGCTCCTCGATGAACAAGTCGATGAGCTCGT
CGTCGCCGGCGACGTCGTTGTTGGACAGTTGGTCCTCGAGGTAGGCGAGGGTTTCCTGGGTGAGGTTGGTGAACAATGGA
TTCATGGTGTGCTCCTGGTGGTAGACGAGGAGCGCCCTTGAGGGAGCCCCTCAAGGGAAGTGGTTGGTGCGGTCAGGCTT
GGTGTTCGTGCGTCAGGCGTTTACGCCATTCCTGGTTGATCTCATCGTGAGCCAGGTCCAGGAACTCCAGCGGACAGCCG
TAGTAGAACGGGCCGGCGCTTTCCGGAATGGTCTTGTAGCCCCATTGCCCGGCCGAGCAGTCCAGCAAGTGCAGCTCGAT
CATGGTGATCGAGTGACCGGCCGGCTTGCCGAAATAGAATCCGGCAAGGTGAAAAGTTCGTTTTACGACAGTCCAAAGTT
CATTGCCGGAGACTGCGTGTGCCAGCACCTCGGTGCTGCCTGCGAAGGTACTGGTTGGGGCCAGCAGCTCACGCAGCAGG
TCTTCCTTCGTCTGATGTGAGAAAAGCCATCCCATAGGTGTGCTCCTGTGGTTGGCCAGGTCTCCCTGACGGGAGAGCCC
GGCAGGGGGTGGTAGGGTGTTGCTTCAGCGTGTTGCGCGCTCAAGGCGCTGGAGGAGCCGGTAAGCTGTTTGCCTGGCCT
GCCAGTCGCAGTACAACCAGGGCTCGTTGTCGGCCGCATGCTCCGGCGAAGGGGTCTTCACCCGCTCACGCAGTGCGATG
ATGGCCTGGTCGACGACGGATCGGTCGAACACTGCCGCGCGGAACTGGTCAGCCAGATTGTCCAGTCCGACCTCGCGCAG
CGAGGCGCAAACCAGCAACAACTCCTCGCTGAAGCCTGGGGTGTTGTAGATCGATTTCATGCGGCCTCCTTGCGGGTCCA
CACCTGTTGTTTGAAGTCGAGCTCATAGCCCAGCGCGCCGAGTCGCACGCACTGTTGGCGTAGTCGCTTGCGGTCGACCG
TGGTGTCCAGCTTGACGATTTCGCCTAGCGGCCAGACCGTCCCGAACAGTTCTGCATCCGCGTCTGTGTCCAGAGAGGCG
GCCTGCTCTGCAGGAGCCGGAGTGTCCATACCGAAAGGCGCGTCCCTGGTTGCAGCGAGGCACAGAGGCTTCGACTTCTG
GGGTGACGGCGACTTTGGTGGTATCGCCGGTACTACTGTGGGGAGCTGGATCTCGGGCACTTCGTCGAGCGGATCCACCT
CATTGCCGGAAAGACGCTCTACCTCCTCGTCGCTGAGGCTGTCCATATCGTTCAGCGTCATGCTGTCCACTATGGCGCGG
ATCTCGATGACTAGCCGACCGCCGGCGGAGTAGTGGTGGGGGCGGATATTGGCGATGACGAAATCACCGCGGTACTTGCC
TTCGGTGTGTTGATCCAGGAAGGCATCCTTGATGATGAACTCCCCGATCGAAGTCCAGAGTTTTCCCACATTGAACTCGC
CATAGCGACCGTTGATGGTTCGAATGGCCAACTGTCCGGGAACTGTGATCATGGTATGACCTGCGAAAAGGATGATCCCG
CCGAAGCGGGATCGTGGTGAGGGGATGAACTACTCGGTAACCGGGTGTTCGGCGAGGCACTCTTCCTGCAACTCGATGCG
CAGGTAGGAAGGCTCATCGGACGGCTCGACTATCGAGAGAACCAGCAGCGCGAGCGCGCGGTCCTGCCCCTCTTCCTCCT
TGTGGTACAGGTGGAATTGGATGTGCTTCGCGGTCGGCTGGAGAGAAAGCTCCTGCCAGGCGGTTTGAACGACACTGGCC
AGCCGGGTGCCCATCTCGTCGACCCTGGCAGGGTTGGCGAGATAGACCAGGCGCAGCCAGGCCGCGTTGGTGAGAACGAT
CGGGAGTTTGAGGTACGCAGCAGTGTGACCTGCGCTGGAGTGGTTGTTGAGAGAGGTCATGAGAGGTCTCCGAATACGCA
GGCCTCTCCCGCGGGGATGAAGTCCCCGCGTGGGTGGGTAGGATGCACTCCATCACTTGCGTGATGCGACCGCGCTGCCG
AGTGCTGTGGCGGTCTTGGGTGGGTTAGCGCTTGCGCGCCGGAGCCTTAAACCCTGGCTCCTTGGGAGGCGCTGTCAACG
ACAGCGAGCCAATTCGAATTCGGGGGCTGGCGGGATTTCCCGCCGGCCGGATTGCTCAATCGTCATAAAGCGGCAGGCCC
TTGAGCGTGGGCGCGAATGGATCAAGTACCAGGAGCCGAATGTCTGCTTGTCCCGCGAGGTGAAGCAGATCGACAAGGTG
TTCGTCGAGGCCGAAGGCAAGCTGTTCTTGTCGAAGAGCGGCTGCGGTCATCTCCAGGATCTCCAGACACTCATCCGTCC
AAGGGGTGTTGTGCAAGCGAAAGCCCACGGCACCGTGCTCGTTACACCAGAAGAGCGATACCAACACCGAGGGCAAGCGG
ATTGGCATCTTGGTCGAGAACCTTTCGTAGAGGCTATTCCACGCGCTGATGGGTAGATGTGTTTTTCTGATCTCCCAGCA
GGGGTTATAACCGCCGCCGAACGAAAGGTATCGAACGACACTCTCGGCTTCGGTCAGTGAGTATTGATCTCCGATAAGCT
TTTTTCTGCCGTCCTGCTCACCGTAGATGCTGTAAAGTACGTTTCTCACTATTCCTTGGTGAGGGCATTTTTCAACTCTT
TCATCGTCTAACTCTTGGACTACCACGGCGAAATCGTTATTGAAGTAGCACAGCTCGCACTGAAAATATTCGTCTGGGAG
GTGCTGCTGGGTTTCATGCAGAGCGCGGTAGCACGGTAGATTGTTGCCTTCGTCAGTGATCTGCAGGTTTCGCTCGATGC
GGAAACTCTCATAGCCGCGAATAAACGGGTTGAGTTGAGTCATGGGATATCTCCGGTTGCCGGGAAGCCCATGCCCTTGG
AGGCTATGGACCTCCAGGTGAGTAAGGGGCGCCGAAGCGCCCAGGTGATGTCGAGTTGTTGCAGTTAACGCGATAGGTGT
TCTTCTTTACGCCGAGCCTTCTGGTATCGCTCTACCAGGCCATAGTATTGAAGTACCGGTGATGTGCTTGAGTCGCAACC
ATGCTCACAGTCGGCCTGCGCCAAAGCTTCGATTGCATCTTGAATTACATAGAGCTCGATACTGAGTCCGGCATCGCCAA
ATTCATCAACCCGGCCTTTGCACTGGATGTACTCGAACCTGCTTTGTGCAGCATGCACTCCAATGTTGAGTAGAGCTGGG
AGCCTTTCTGGATATGTGTTCATGGTGAGTTCCAATTCGTAATCAAATCGGAGAGGTCACCTCCCAGCAGGGGTGATGAC
CCCGTCTGGGTTGAATTAGAAGATGCACATGGGCAAAAGGGTAAAGCTTAATATTCCAGTTGCGTCCAAGTGCCGTGGTC
TAATGCCTGTTGGGCAAGTTCGCGTGAGGGGTAATATTCAACCGACTCGCGCGAGGCTGGTCCTTCTTCGTCAGAAGTGC

Figure 2

FIGURE 2I

CGATGTAGTAAGCACGGTTCTTATTCATAAGTACCTGCAACGGCAGTTCCAGGCTGAGATGGGCGAGCGCCAGCTTTCCG
TATTGCATGGTGATTCTCCTTTCGGTAGGAAACACCAGTCCCATGGGGACTGAGGTCCCCTAAATGGGAAATGTTTGGCG
ATTATGCCAGGGCAGATACCTGGGCATTGTTGTTGGCCAGGAACAGCCAGGCGAAGTTGCAGAACTGACTGCTGAGCAGC
AGCTCGTCCTCGACGTCGCTGAGGCTAACCTCCTCAACTTCCGGCAGGGCGGGCAGTTCTACGATTTGGATGTTGCTCAT
GATGATCTCCAGATAATGGAGAACACCACGCCCTGGGGGGAGATGGTGCCCCCCGGATGGGTATGCGATGAAGAGTGGCG
CATAGCGTAGGTAGTGAAAGGCCCTCGGGATGAGGGCCGGTATTCAATCAAGCCGTGTTCAAGCAGCGGCTTGTTCAGCG
TGCTCAGCGGGGGCTGCGGGCTCACCTTGAGGTTCGCCTTTGCCTTGCTGGGCTTTTTCAGCTTCTTCCTTCGCATCGTA
GACGGTGGTGCCATCCACTTTGATCCAGGAGATGAAGAGCAGCCGGCCTTTGAGGCTTGCGTCGGGCTTGCCTTGTTTCT
CGCCTTTCTGGTGGATGAAGGGATCCGCCCAGATATCGCCGATACGGAAGGAAATCAGAACCTTCTTCTTGGCCTCGACT
GCTTCCTTGCAACGGCGGACAAGCTTTTCAGCCTGGGCGCCGACCACTTTGCAGTCGATGTAGGAGTATTCCACGCTGTC
TGCCGCGCCATGGAGGGCTGCGACGGTTACGGCGAGGAATGGTTCACCTCGGCGGATCGGTACCTCGCGGATGCGATTGA
GGTAGCCGATACCGGTGGTGTGCAGGTCGAAATACTTGGCTTCTTGGGCTTGGGTGTTGTTGCTCATGATGATTCTCCAG
TTTCGTAAGCGAAAGCGGAGAAACACCTTGCCCGTCAGGGAAAAGTATTCCCCGCATGGGTCGGTATTGGGAGGTAAATC
AGAGGGAGAGGGCTTCCATCAACCTGTGGGTTGATTGATCGCGCTACCGGAAGCTGGTGCGATCTTGGGTTCGATCTACA
CGGAAGTCCGTGTCGTAGTCTTTCAGATCATGACTACTGGGGGAGGGTTGCCGACAACGTCAGCAGCTCACGTGGTAAAG
GATGGGCAGGGGAAAAATGAATGTCAATACCCATGAAATGAAAAAAGCGCCCAGGGTTGGGCGCTTTCATGGTCAATGAT
GACGGGGAAGAATCCGATGTTTTTGGCACCAGGTCTTGACCTGAGGCCAACTGAACCAGAGGGCTATGCCGGAGCAGATA
ACCACCAGGCTGGTCATGATGATCATGCCGAACTCGGTCTGGGACATTTCACTCCTCCTCGGTGAGAGCGATACCTATGA
GCCAGATTACCATACCCATGACGAAAGCATATGCCGTCAGCAAGCGGTGCTCGCTGAGGAAGATGCCCACCAGGCTCGCC
GCAACGGCAGTGGTACCGATTTTTCGGAGATCGTCTGCGAATGCAGAACGTTTCGATTTCATGGTTTTCCTCGAAAAATG
CCGGAAAACCACCCTGCGGGATGGTTCCCGGCAGGGTTGGGGGTGGTAGCAGTGGCTATTAGCGGCCGAGAGCAGCTACT
TGTTCATCCGTCATGGTGGACAAGTAGTCCTGCCAGCGTTGTTGGTATTGCTCCTGAGTACGCCGATGGAGTTCCCAAGA
GACTGCGCTACCAGGAGGTGCTGGCAAAAGAGCAATCTCATGATCGTTCAGGGACTCTAACTGGAGAGGGGGCCTCCCCA
TTTCGCTTTCGGAGTTGGGTGCAATCAAATGGTCGAGGGAACCCATCTTGAGCTCCTTCCTGACTGGCGTTTCAGCTTTG
GCGGTCGCCTTGGCCGGCTTGGGTTTGCGTGCGCGCTTGGGCTTTTGCACCTCGCGCGTATTGGCTAGATACCAGTCATT
CAGGTTGACCTTCTTCAGCCGGACGATGGCTGGGATGTCGCTGATGTGTTCTAGCTCGCCAGCGGTCGGGCCGGCCAGGA
AGGCAGGATCGATGCGCTGGACTCTCAACGCAACCGGCGCTTGTGCTTGCACTACGACCGGAGGGCGACTGAGCGCTCTG
CGGAACTCACGCACAACACTGCCGAGGGCAAAGCCGATTTCGAAGCTGGTGGATTGCTTGGACATGGTGTTTCTCCTGCA
GGTTGAGGGTGCGGAGAAAACACTGGCCCTGGCGGGGAAGTGGATTCCCCGCTGGGTTGCGGAAGGGAGCTTCTATCGAC
CGAGGCCGACCGACCGCGCTACCAGAAGCTATGCGGTCTTGGGTTGGATCAACACGGTGGGCTGTGCTGCAGCATTTCGG
ATCCGAGCTGCATGGGGGCATGTCGCTGACAACGTCAGCAGATCACGCGTTCATCATCGAAGCGGGGGGCGAAAGTGTCA
AGGTCAAGCAGGAGGGGGAGCGGATGCAGCGGGCGGAGAGCGAGGGCGGAAGAGCGCGGCTGGTAGAAACCCTGTCAAAG
GTTTCCCCAGCCGTGTCTGGAAGGGGAGTCAAGTGAGCCACCTGTGGTTGAATCGTCGATCCCTGGGCATTGATCGTCTC
GATCCCATCACCCGGCCATTATCGTGGCTTGGCCAGCAAACAGTAGGCACGCATCCGCGTACAAAGGGAGCCCTGCGTAT
CACCGGCGGGCCACCGGCAGGGAGAAGGATCCCGATGGGTAGCCTGATAGTCCTGGAGCAGGAGCATCAGGCTACCCATG
GAGAGGGGAAAAGGAGGGGCCGTAACACCAGTACGACCCTTAAATCGAGGAAACACCGAACCTCTTGAGAACCCTATGTG
CCTTTAAACTCGCTCGACACGAGTTGACGGCTGCATGACAAGCAAAAGAACAGGGAGACACAGGGAGAATTCTCAACCAG
ATCCGCACGCCCCACCCGGTGTTGTAGCGTGCTTACGGTTTGGCGATGAGGCGAACCTGCATCACCAGGGCGCTACTGAT
GACCGTCAGTAGCCAGCGGGGTGAAATCAAAGTTGATGGCATTATGTGTTCAGAAGGCCGGTTTATCAAGCTGCGCCGGG
GTTTACTGGCCTGTGGAGGTTATTTGACCGCGACCACCGCGCCCTCGGGGATGGGCGTAGGGAGCGCAATGACAGTGGCT
CCGAGGCGGCGCGTGACTAACTGGCCGTTGCGCCTGACGCCAACCAGGTCGGATTCCGTACCAGCATCCACGGCCGCCGC
CAGCCCCGCCGGCGGCGAGGCGATGAAGGCGTAGCTGGTGCCGGCGGCGATGTAGCCCTGAAGCCGCGCTGGCTTGCGGA
ACCAGGCCGGTAAACCAAGAGCGGAGTCCGCCGGCGAACCGGTGAAACCGGGGTTGGCGTGTGCGTACTCCGCCAGACTG
GACCGGAAGAGCAGCAAGCTCCTGGCCAGGGTGTCGACCTCGGCCAGTTCGGCGCTCGAGGTCGCGTGGTCGCTTTGTAC
ACTCAGCCAGGTCCCGGTGATGAGCGCGAGCACCAGGACGATCCACATCAGCGGCATACGGTAGTCCTTGAGTCGTGGCG
ACCTGCTAGTTCACGCAGGTAACGGTCACGTTGGTCGAGTCGGGGGTCTTGTAGGGTTCCGGGTAGTTGAAGCCTGGGCC
GATCTCAGTGATTAGCCCGAACGAGTAGTTGGTGCCAAATGTCTGGTTGAAGCCCACTGCGCAGAATCGGCCGCCGCAGG
TGATCGCGTAGCGGACGTAGTCGTAGTTCACCATGGCGCTATCCTGACCATGCAACGTACACCAGTCCTTGATCACGCCG
GCCGTGGTGTTCAGGCGCATGGAGGCGGCGGTTCGTCCCCATTTCCCGTTTTGGCAGGACAGCCAGAGTCCGGTGCTGGT
GATGCCTGCCAGCCCGTTCGGCGAGCAGTTGGCTCCTTCGGTGGCCACGCCTTTGAGCTGTAGGTACTCGCCGACTTCCG
TCCGGCCCTCGGCGGTGAGTTTGCCCGCTTTCATCTCGCCGCCGGTGTAGACGTTCTTGTTCATCCAGGAGCGCACCCAG
GTGCTGTCGCTCATGTACCAGCCGCCGCCCCATTTCTCGTTGTACCAGCCCGTGTCACCACGGGTCCTGAACCAGCCTCC
GGTGTACGTCTCGCCGGCGACGTCAGCAGTGCCGGCTTTCACCGTACCAGTGGCTGTCACGTTGCGCGCGCTGATGTCCG
CGCTGGTGGTGATGTTGCCGCTGGCCGTGATCGCCCCGGCTGCGGCGATATTGTTGCCTCCCATGTCCAGCGTGGTATTC
ATCCGGTTGAGTTCAGGATGACCCGGGACGGCATTGCGGTAGAGGTACTCGTTGGCGATGGCGCCGTCCTGGAAGAACAG
CGCCGTCGCCAGATGTCCAGCGCCGGGAGCGCTACCGAAATTGCTTAAGGCCACCTGCCAGGCGGCGCCCTGGGCGATAC
TGGTGTTGGTTTTCGAGATGTAGCCCCCGGTGGCTCCCATGAGCTGCGCGATCCGGCGGATCGAGAGTTCGGAAGCTACC
TGTCCACCCGTGGTCACGATCAGCGTTTCGAGCTGGTTGGCCGCCGGCTTGCGGGCCAGGACCTGGTATTGCTGGCCGTA
GATGTTGGTGTCGCGGAAGCCTGCGGGCAGGTAACGGGTGTTGCGCAGCATCGGGACGGTGATCACCGCCGGGGCCGTGG
CGCCGGCGCTGGCCAGAACCGTGCTGAAGTTGTCCTTCAGGTACTTCTCCGCTGCCTTGGCCACCTGCTGTTGCTGCTGG
GCCGCGTGCTGCTCGTCCAAGCCGTCCAGGTAGCTCATCAGGACCGATATGCCACCGGCGGTCGCGATGGCGATCACGAC
GAGGGCGATCATCAGTTCGATCGAGATGAATCCACTGCTGCGCGTACTCCTCATGGTGTTATCTCCATAGGCTTTGCCGG
GACTGGCGAGCTTGGTCGTCCTGCAGCAGATGAGCCTCCTTTCGGCGGATCTTGTCGCGCATGGACGGGCCGTCGTCGCC
AGTGAACAGCAGTGGCTCGGCGGTCAGGCGGCGCGAGGAACCATCGCTCTCCAGCGCTTGGCAGATCACAGCTTGGATGC
CTTCGGCAACCACGTCGTAGGCGTTTCTCGCCAGTGGCTGGGCGAGCGTCACCAAGCGCTCGATCGCCTGATGACAGCTT
TTCGCGTGAATGGTGGCGATCACCAGGCTGCCGGTCAGAGAGGCCTTGCAGGCCTGGTAGGCGGTGTCCTCGTCGCGGAT
CTCGCCAATCAGCACCAGGTCGGCGCCGGCCCGCAGCGTGCGCAGCAGGGCCTCGCTGTATCCGCCTGAGCGCCGTGAGG
TTCTGACCTGGATGCAGCGGCCGAGCCCATGCTGACCGCTGAGGTTGGTTTCCTGCGGGTCCTCGACGGCGCAGCCCACC
CCGCCCAACTCCTGCAGGCGGGCCAGGAGCAGAGAGGCGGCGGAGCTGGTCTTGCCTGTCGCCATCTCGCCGCAGAACAG
GACCAGGCCCTGCAACTGCGGATCCATCAGTGCGGAAACCACTTCGCTCGGATAGCCGATCTCTTGGAACTCCCGCAATT
GGGCGCTCGACCGGCGCAGCACGAAGACGCTGCCACTGAAAGCGTCTTCGAGGAGGGTGACGCGAAGCAGAACGCCGTCG

Figure 2

FIGURE 2J

ACGACGATGGCGAACTCAGGGTCCTGCAGCTCCTCCAGCGTTTGCCTGCAGATCTGGAGCAGTTCCTGGACGTGCGACTC
CCACTCGCGAGGCACTTCGACTCGGCGCGCGCCGGCGCCGGCCAGGCCTTTCACGTCGGCGAAGCCTTCACCGAGGTAGA
GGTCGACGAACTGAGCCTGGATGATCGGGTTCACACTCACGAACTATAGGTCCAGGTAATGCTGTTGCTGTCGCTGCTGC
AGGCCTGGGTCGCGGCTGCGGTGGTCACTTCTCCGGTGATCGAGGATCCGCTGTTGATTTTGGTCTGCTCGAAGGTGTTC
TTCGCGATCTTGGTGGCCAGCGTGATACAGGCGTCCTGGGGCAACTTGCTGGTGGTGATCGAGAAGCCCATGCCGGTGGA
CGAGACAGTGACCGATCCGCCGTAGACGTTGTAGACGACGCCGGAGGAGACACTCATGTTCTTCGGCACGCCGTTGATTG
CGATCAGGCTGGGGATCAGGTTGGTACCGCTGGAGCCGTAGCCCGAAGAGGTCTTCAGGGCGCGTGCGTTGGCCGCAATG
ACGCTGATGTTGCGTTGTTCCTCGTTGGCGTTGGACGAACTGAACATTCCAGCCGCTGCGGCCAGGCCCAGGCCGACCCC
GATGGCGATGATGATCAGCACGATGATCATCTCGATGGAAACGAAACCGCCCTGCGTCGGACGGGAAGTGCGTTGCGTAG
TGCTCATGGAGCGCTCCTTGGTGAGGTTTCAGTGGTTCATGGATTGGATGAGCTGCTGTGCCTGGTAGGCGCCGAGCAGG
ACCAGGATCATCAGCGCGCCGATGAAGATCAGGGCGAAGTTCTTCACCAGCCCGGCGGCCAGCTCGATCTGCTTGAGGCT
GGTCTCCTGCCAGCGGCGGCTGAACTTGACCAGCGCCTCGGAGAAGCCTCCCCGGTTGGCGAGGATGCACAGGTACTGGA
TGGCCTGCCGGTCGGGGAAATCGTGACCGGCGCTGCGAAGGGCAACACCCAAGTTCTGGCCCAGGCCCACGCCGTAGCGG
GCAGCTTCCAAGCGCTGCTTCAGCCAGGGCGGGGAGATCTTGATCATGCTGGCCAGGCTGTCGTAGGGGCGTATGCCGGC
GTTGAGCATGACCGCCATGTTCAGCAGGAAGGTGGTGCCCTGGAGCATGCGGTAGATGGACCAGGGCGGCAGCGTCCGGT
CCAGCCAGACCCGGCCTTTCCAGCGGTAGGTCGGCAACGTGACGATGACCACCACCGTGAGGGTGATGACGGCGACCAGA
ACGTAGATACCAGGTCCTGTGACGAAGCTGGCAATGGCGTTGAGCGTGGCGAGCGGGCCGGTCCAGGTGACTGGGTCGGA
GAGCCTGGCCAGGCTGGGGACCATGCGATAGGCCACGATGCACAGCAGGAACACCATCATGGCGGACAGCGCCGAGGGGT
AGAGCAGCGCCTGCCAGATGGTGGCGCGGATCCTGGCCTGGGCCTCGACCAGGGAGACGGCATCGCCCATGGCCTGAACC
AGGTTGCCGCTCATCTCGCCGGCCTCGATCAACGCTCGCTCCTGGGCGGGGAGATAGAGCGCCATGGCGGTGGCCAGACG
CTTGCCGTTGGACAGCCCCATCAGCGCTTCGCGACTGGCGATGGCCACCGGATGAAACGGATGCTGCCCCTCATGAGCGA
AGATCTTGTGCACCTCTGCCACAGCATCCTTCAACGGGACCCCGTTTTCGAGCAGGGTGGACATGCTTTCGTAGAACTGC
AGGCGTTCCTTGCGGCCGAACTGCTTGCTGTAGAAGGCGAACTGCAACTGCTCCCAGAAGCCCCCCATCAGCACGCCTCC
TGCGAGTAGAAGCTGTCGTCGAGCAGATGCTCATCGAAGTCGAGTGGCCCAATGAAATCCTCGACCATCTGCGGGTCGAA
CATGCCCTCGTTGATGCGGCGGATGGCGTGGGCGTGCTTGGTGATGCCCTGCATGGTCTTGACCCAGTAGTTGCGTGCCT
CGGCTGGGCCGCCTTTGGCGAACACACGCATGAAGGCGAGGGTGGGCAGAACCACCTCGGCGACGATCGAGCGGCCGTTG
ACCCCGGAGCCACGGCAGGCCTGGCAGCCAGGCCCCTTGACGTGAACCTGGGAAACATCGGTCAAGCGTCGGACCCGTTC
GACCAAGTCGGGCGCGAGTTGGTCTTGGTGGTCTTGGAAGCGCACTTTGCAGTGGGGGCAGAGCTTGGGCAGCAGGCTCT
GGTTGATCAGGCCGGTCAGCAGGGCCGGATCGAACAGCAAGCCGGGGTCGACGCCCAGGTCCTTCAGGCGCTGGACAATG
CCGATCGCGCTGTTGGTGTGCAGGGTCGACCATAGGCCGTGCCCGGTCATCGCACCACGGAACGCGGCGACAGCGGCGAA
GAGGTCGCGTACTTCGCCGATCATCATGTAGTCCGGATCCAGGCGCATGCCGTTGGCGATGCCCGCGGCCCAGGCCTGGC
GTTCTGCGTCTGGGTCGGTGGCGTCGTAGACCAGTGGGGTCTGGTTGATGCCTTCGCCGCGAATGCGGTATTCCGGCGGA
TCCTCGATGGTCAGGATGTGCTTGGATCCGCCATGGAGCTTGTCGAGGCCTTCCAGGGTGACCTTCAAGGTCATCGACTT
TCCTGACCCGGTGGGGCCGGACAGGATGTTGATGCCGTAGGGCATACGCATCATGCGATCGAACAGTGCGTTCTGCTCGG
GCAGGTAGCCGAGCTGCTCCAGGCTGTCGAGGCCGGTGTCGTCGTAGAGCAGTCGCAGGATCATCAGGAACCCCCCGGCA
CGCGGGCGGGTGGCGATCCGGGCACTGAACAGGTTGAGCTTCTCGACGAAGGTCTGGCTCATCCGCGCGTCCTGGTCCAG
TTGCGGCTTGAACAGTGGCTCGGCCACGTCGCACATGGATTGGTAGATGGTTGCACAGAGTTCGTGCAGCTCCTGGCTGC
GGAACTGCTCGACGGTCTTCAGCAGGCCGTCGACGCGGAAACGGATCTTGCTGCCGGTGCCGGCGGGACTCACGACGAAA
TGCACGTCACTGGCGCGCAGCTCAGTGGCCTTACGAATGATCTTGACCACGCGCTCCTGCATCTGGGTGGCGCGCTGGCC
ATCGCTATCGACCAGGCCGTCCATGGCGACGGCGCGGTAAAGCTGGTGAATGGTCTGCAGGTCGGTGAGGTTGAGCTGGT
ATCGGAAGCCACGGCGATCGAGACGGTCGATGAACGACAGAACGTGTATGTCGTTCTGGTGGCTCTCCGATAGATAGAGC
GTGCCGTCGGCAGCCAGAGCCATGATTTCCTGCAGGTGCTTCGATACCTCCCATTCACCACCGTCGGCGGTGAGCAGTTG
GGTCACCATGGAGGGCTGCGCAAGCGCGGCAATCTGAAGGTTCGTCATGGAGGTGCCTACTGAATGAACGGCTGCGGTAC
AGCCCCGGGCAGCGCCGGCGGAACCGAGGCGCCTTGGGCCGTAGAGGAGGCTTGGGTGGGGGCAACGCTGGAGAAGCCCA
CGGGCACGCGGTTGCCGTCCTTGTCGGTGAGCACGACCTGGTCCAGCGAGATTGACTCGACGCGGTATTTGCCAGGCAGC
TCCGCGCCGCTGGCGGCGTCAACCTCGTACCCGCCCGGGAACAAGAACGTGGCAGTCATCTTGCCGCCGGCGCCGTAGAT
GGTCCGCACCACTGGCGGAACCGGCTGCGTGATCGTCGGCGGCGCCGCCGGCACCGGCATGGCGAACGTCTGCCCCTGGG
CGCTGACGACCTGGCCGGTTTCGCCTTTGCCTTCCAATTGCCGCTGCGCCGTGGCCAGGCGCACCTTGGCCTCGGTGAGG
ATGGCCTGGGCCTGGATCTCCGCAAGTTCGCCAACGGTGCCAGCGCAGGCCTGGCCAGACGCGAGGAGGAAGAGCACGGC
CACGGCCATGCCGATCGGCTCAGTTCGCATAGATTTCTCCTGTGACAGTCCAGTCCAACTGGCTGTCCTTGAGCGTGGTT
TCGAGGTTGGTGATGCGGACACCGGCAGCGGGCAGCCCCTGGAAGGTCAGGTCTGCCGGGAGCCGGGTCTGGGCGCTGAA
GGTGAATTTCTTCCAGGAAGGCAACACCACTTGCTGTTCAGTCGCAGCTTCCGCGCCAGGGAGGGGCGGAGTTGTCTCCT
GGCTGATCGACAGCTTGGGCTCGACCCCTTGACGGTACAGGTGGCTCGTCAGCGCCTGCAGAACGTCGTCCGCCGGCAGT
AGCGGCTCATCACTGCCGATGGCAACCTTCAGATCGACCTTCAGGGCCGCGGTGTTGCCGTTGTCGATGACGAAGGCGGG
GCGGTCGGCGAACAGGTGCTGGCTGGCCGCTGTCAGGTCGGCTGCTGTGCTGTTGCCGGTACGGTGGTAGGTGGCGACCA
GGACCCTGCCGTCGCACTTGCTGGATTCGAAGAGCCAGCCCTGGATCGACAGCGACAGTACCCCCGTTGCCTTGCTACAG
GCGCGTAGCATGTCCTCGAGGTCAGGTATGAGCGTCCAAGGCTTCTGCAATGACGCCAGGTCCAGCGGCTGCTTGGCCTG
GGCGTTCTTCTCGGCCAGCTCGGCGAGGCGCCTCTGCTCCTCGAGGAGCGCGGCTTGCCTGGCGAGCTCTTCCTGGTGGG
CATTCCATAGGTAGTAGGCGGTTAGCGACCCACCTACCACGCAACCGAGCAGGGCCACTGCCGTCCACTCCCTGGCGGAC
AAGCCGAAGGTGAGTTGCCGGAGGCGGTAGTCGCGCCGCAGGCGCTTCGGCGCGAGCAGTTCCTCGATGTCGAAGTCCTT
GACGGGGAAATCGAACCCTTCGGGAACGAAGACCTGTGCGTTTCGCAGCACGCCGCGCGTAGAGAGCTTCCTGACCCGGT
CCCGGGCCTCGTCGAGGGTGGTGACCAGATCCTGCCCGGGGACAATCGCGCCATCGAGCGTGGCGACCAGCGCGTAGCGG
TCCTCGTCGACTTTCCAGCAGGCCAGGAAGTCGCCGTCGAACTGGCCTGAAAGCGCCGAGGCCAGGGAGTACATCCCCTT
GACTGCGCCTTGCGACTTCGAAACGAAGCCGGCCTGGATCACCGTCGGTGAATGGCGGATGGCAACGATGTCCAGATGCT
CTTCCTTGCCCAGCTTGCGCGCTTCCTTCATGTACTGCCGCTGGCTGGACAGCGGCCGCCAGAACAGGCCGCTGACGAAC
TTGTTGCCGTGGTAGCTCAGGATCGAGACGTCTGGTCCACGGCTGCCGAGGTCAGGCTTCTCCATGGTTCAGCCCAGCAC
GACGGGGGTGATCAGCACCACGATGACCTCGCGCTTGGTATTGCGGGTCAGCCCGCCGCCAAGACCGAAGAACCCAGCGT
CGCCGGTGCCGACCTTGTTGGTGTCCTCGGTGGTCTGGTCGAAGCCGGAGAGTACCAAGGTCTCGCCGCTGCGCAGACGT
ACCTTCTGGTCGAACAGTTGTATGTCGTAGCTCGGGAACTGGGCTTTGGAGTCCCCGCTGGTCTGCATTTCGAACGTCGG
CCGGGAGGTCATGTTGATGTTGATCTTCAGCAGCATCTCGCCGCTTTCCATCACGAACGGCAGCAGGCTCATGTTGTAGC
CGCTGGTCACGGCGCCCGGGATCAGCGAGGTGGTACTGCCGACCTGGGCGACGTTGGAGATCTGGCTGGAGGCCAGGTAG

Figure 2

CTGTCGTAGCGGCCGATCTGGATCGGCGCCGACTGGAGGTTGAGCGTGGTCACGGACGGGGATCGGACGGTCGAGACGCG
GCCCTGCTGGGCCAGCGCCTGGACCATGGCCTTGGATCCTGCCCAGGCGCTGTTGGCGGTATCCAGGATGCTCACGGAGC
CGGAGATCGCGCTTTGATCGATGCCCGGCATGGTGTTCTTCAGGCCGATGCCCCACTTGTTGTTGAGCGACTTGTAGACC
AGGTTCCAGTCGATCCCCAGTTGATCCTTGTCGGTCAGGGCGACCGAGAGCACGTTGACGTTCAGCAGCACCTGCTTGGT
GATGCTCTCGTTCTCTCGGTTGACCAACTGCTGGACACGGTTGAGGACTTCTGGACGGTCGGTGACGGTCAGGGTGCCCG
TGGCACGCGACAGCGACATGCGTCCCATGCTCGGCGTCAGCATCGAGTTGATGCTGTTCTCGATGTCGCTGAGGATCGAT
GTCTTCAGCTCCGAGCTGGTGGTCTGCTTGCTGCCGGAGTCGCCGCTGATGCCGGAGCTGCCATTCTGTCCGGTGGATCC
GGAGCCGTCCCCGCTGATGCCGGCGGCCGTCGTCATACCGGAACGCACGGTGGAGTCCACCGTGTTGACGTCGTCGAAGG
CGTACATGCGGAAGGTCCGAGTGTCCAGGTAGTAGAACTCGACCCTTTTCTCGGTTGGGTTGTAGCGCCAGGACACGCCG
GCTCGGGCGGCGATGAGATCGAGGAACCCGCTGACTTTGCCGTTCCATTTGATGTTGTACAGAGACCGCGGACCGTAGGA
CCCGAAGCTCGACCCCATGCTGCCGCCGGCACCGAGCGACATGCCGTTGGCGACGGAGGCAGGAAACAGCATGGTGGCCA
TGTCCTGGCCGCCTTGGATGGGCGGCGGGGCGTTGCTCGCGCGCTGCTGAGGTTGCACGGCGAAGGCGGCCGGGTTCAGC
GCGTCGGGCGTGATACTGACCGCCATGTGGCATTGGTTGATGACTTCCTGGGCGGCCTCCTGCAGCGACGCTGCGCCTGC
AGGGCGCCACGTCACGATGCAGTCACTGGACAAGGTGTGCGAAACGCTTAGGGGTTTCGTGCTGACCCAGGGTTTGTCGG
AGAACACCACCGTATCGCGCCGATCCGGTCGGGTGTTGCGCACCTGCGCTGCGATCGTAGACGCGGAATCTGCCGTAGCT
TCGCACACGATCCGCCGACTCGTTCACCCGCGTGACCGAGCAACTGGCCAGCGCGGGGATCAGGAGGCAAGCCAAGGAGG
TTCTTCAGGGGGGCACGCATCAGTTCTTGCGCTCCTTGATGATGATCAGGGACTGCGGGCGGCTGGCGTTCACCAGGAAG
GGCCGTTCGGCAGCGTCATACAGGGGGAACAGCTCGGATACCGCGTCCTCGAAGGAGCCGTGGAAGGTCAGTGGAGCCTC
GATCGGATAGTTGAGATCCTGCGGCTCCCAGCGGACGGTCCAGCGTGCGCGCTTTGCCCAAGCTTCCAAGGTGTCGCGCA
GGGTTGATCCCGTCTCGGCTGACCAGGCCTGTGCCAGCGGCCGGCTCGGTTCCGGCGGCGTGAGCACCTTGACCGGTGGG
GCGGAAGCCACGGTGGGTGGAGTGGGCGGCGTGGACTTCACGGCGGAGGCCGCAGGCTTGGTCGGTGCCGAAGCCGCGGG
AGGAGAAGAGGGGTGGCCATCCTTGGCCGGTGCGGCTGGGGGGAGCACAGTGGTGGATTCAGGCTTCTTGGCCGGAGCGT
GGCTGGTTGTGATCGGCGCCGGTGTCGGCACCGAGGCGACCACGATGGGCGACTCCAGCGTGCTGACTTTCTCCGTGGAG
GAGGGTTGCGCTACCGCCGGCGGAGTTGGGGCAGCGGGCTTCGCATACAGTTGCTGGACCGGTTTCGGCCTCGGCGCCGG
GGGAAGTTGATAGCCCGGGCGCAGCACGAAGCAGACCTGCCGCGCGACCTCGTCGACCTTAACCTGCCAGGCTGGGCCGG
AGAGGACCTGGAGGGTGTTGCGCAGGGTCATCGGGCCGAGCTTGTACTGAGCTGCCGGCAGCGGCCGGGTGTAGAGGATG
TTCACATGACCGGCGTCTGCCGGGCACAGCGAGTAACCCGAGCGGCTCATCACGTACTGCATGGCGTCCTTGACGCTCGG
GTTCATGCTCGACGGGATGGTTACGTCGATGATCTGGGCCATCAGGTCGCGTTGACCGGCATCAGGCTGGGTGCTGACCA
GGGTGTAGCGCCCATAGCGCACTACGGGCTCTTTCTCCGGCTGCACACCGTTTGGATACAGGTCCGGAGAAAGCAGGTCT
GGATTGCGGCTGCTGTCGAAGGCTGGCCTGGCCGGCTCAGCCTTGTGGGTCGTGCAGCTGGCGAGCAGCAGGCACAGCGT
GAGAGTGGTCAACTGCCTGGTCATTGAGGCCCCCTTTGATGATGGGGCACACAGTGGCAGGCCATCCGCCACGCATCAGC
AGGAAATACTCGGCAACAGGGGAGGGGATTTCTCGATGACTAGGCAGAGGGTGAGCGAGGCAGGCGTAAGGGAAACGGAT
TAGGAGCGATGAAGGTTTTTCATGTTACGCCTAAAGCGAGCCAGGTGGGAATCGTTGTTAGAAATGGATAAAATAGAAGT
TCTAATCGTTTTAGAATGCTTTGTTGAACTCAGGTCAGACAGTACGGGGTGCTAATTGAGCTTTAAATACTATTGGGCTA
AATTTTTCTGGGGAGCTTTCTTCTTTGTTTTAGTCGCTTGGAAAGGCTCCGTATTTCCAAGCCTGGCATCAGTTAACCCC
TTGGTAGTGGCTGGATTCAGTACTATCCTGTTTCCTTTCTCGGTAAGGCTTGTTGAAGACTTCGCTTTAAAATATACGGA
AAAAGAGTTCTGGGTCACAGGTTTTTTCTCCGAAACCCCTGCAAAAACAGGATTGTATGCAGTCTTTTATTTGGCTTGTT
ATTTGTTTTCAATTCCCTTGGGGATGATTTTTTTATTCTATAAATACGGAAAGGCCTCGTAGGATTTAACGAGGCCTTGC
TTGGTTTTAAATGGACATTACAAGGTTGTTTGCTTTTTCTAGAAGGCCTTCGTCAATCATCGCCCCGGTAACTGCCATTA
CCAGTGCGAACCCCAGGATGCCTATAGGAGTGGCCGTTGCCGTGGCAAATGCAATACCCACAAGCCAACTGGCGGCCGCA
CCAGCAGCTAGTGTTTCAACTTTACAAAGAATGGTTTCCAGTCCCCGGTTTCCGTAGATATCTTGAACTCTTGGTACAG
GCTGGCGGCGTCAATAGCTTTGCCTACAACTCCAAAGCCTTTGCTAAATTTCTCAAGGCTCTTCGCCATCATCTGCTTGT
CTAGAGAATCAAAGGCTTTGGCAATGGCTTGCCTATCTTTAAGGCTAAGTTTTTTATTTAACGCATCCTTGTGTTTTTCA
AACGAATTGATTGCTTCCGCCGAACTCCTGATATTTTTCCCCCTGGCGCCTTCGGCCAGTTGGTGCGCCATCTCTGATGT
TCGTGCGCCAAATTTCTCAGTTACTTCCTTATAAAAGTCGGCAGTAAATTTAACAGCATCTTGCAACGCTTGTTCCTCGG
CCGCCTTGCGTTCTGCTTCTGTCCTGCGTCGTGTCGTTTCGGCGTTCAGCTTATCTAAGCGGGTCTGTAGACTGGCTATC
TCATCCACAAGTAGCTCACCATTGTAGATGGCTTTCTGCTTTTCTAGAGGGGTTGGGGCATTCTTGATAAGAGTATTGTT
TGTGGTCTGCAGCTGCTTTTTATCAGCAATTGCTTGCGTAACCAGTCTTGTGGCACGTTCAAGCCATTGCTCTGGCGTTG
GACTGACCAATGGAGGCAAAGCACTTCCCGCTTGATCAAGCTCGGATTGTACTTTCATCTCGAGATTCGCCTTTCTGGTT
TTGTAGTCCCGCTCCAGTTGAGAGAGGTCCTCTTTTTCCTTTAGCGCTTTTTGCTTGTTGAGCTCCTGTCTTAGCACTTT
TACCGCTTCAGAGAGTGAATCTTGTTTGTAGTAGGCAACCGCCCGCTCCAGTTCATCTCTAACAAGCTGCTGAAAAGGGT
ACTTCTTTCTTTCTTCCCAAGAGTTCTTCTCCCAGATCTTGACCCAGGCCTCATAGATGGAGAGGCTTGACTGCTTTTCG
TAGTCAATGAATGGGCCTTTGGTAAAGGTCTTTTTGTATCCTTCAAGATCAGATTCCATCTTGTGCAGGACTGCCGACAT
GACATCGACTCGCTCTTGCCTTGTGGCCTCCCTTTTTCCTTGAGTGATCTGCTTTAGCAAGCCATCGACTTGATCAGATA
CTTTCTTGTTGATCTCCTTCTTTTCTTTTGTTCCTTGTTCAACAATGGATCGGACTGGATAATAAATGTTTTTCTGTAGA
TATATGTCTCTTTGAATCAAGTCGCCAAACATATTCCCGGCAACGGCCGCGCTATCGATAGGGGGAACCTCGTATGCGTA
TTGATCGTCTGGACCTTGTGCGACAATAACCATGGAACCAGGTACTTCGTTGTCATTGGACATTTAGACTTCTCCATTGG
TGAGTGTGGTACAGAAACAGTGCCGATTCAGCTCTAGAAATTGAACCCCATGGCTTCAATGTGTGGGCAATGCACTAAGA
AGGCACGTCAATGCGCTGTGATATAGTGGTAATAAGCACGGTTGCTTAAACTCGATATTCAGGAATGGGATTGGCTGAGC
CCACGGGATATTCCCTAAGAGTCATGTATCTCTAATAACGAGTTCATGCGCTCAAGATTCACGCTGGGTATGTGATGCTC
TCTGCAAAATTCAGCTACTTGATCCGCAACACCATTCAAGGGTGAAGATAGGGATAGCAAGCGAAGAGACAGTTTTGTGC
GTTCCATGTTAAACGGGAGATACACGTCGAAGACCTAAGCCATTAAAACGTGCTGACCATCCAGTTGGGGCCCATTAGGA
TCTAGCACAGCACCCATGGTGCACTGGGGCTATTTCTCGGATGTACACCGTTTTGGTACAGGTCCGGAGAGAGCAGGTCT
GGATTGCGGCTGCTGTCGAAGGCTGGCCTGGCCGGCTCGGCCTTGTGGGTCGTGCAGCTGGCGAGCAGCAGGCACAGCGT
GAGAGTGGTCAACTGCCTGGTCATTGAGGCCCCCTTTGATGATAGGCACACAGTGGCAGGCCATTCGCCATGCTTCAGCA
GGAAATACTCGGCGGAGGATCAGGGCTTTTTTATCCAGTTAGGCTTGCGGCTAAAAGCGCGCGGTTTCCGTCGACCTGCT
CGGGCTCGATGAGGTCCTGGGCAACCTGGCGTTTCGCGTTGAGCAGTTGCCACAGTTGCTCGTCGATCGTGGCCTCGACC
AGTGGGATTTTCACCACGACCATGCGGAGCTGGCCGTTTCGGTACGCGCGGTCTTCGGCTTGTTCCTGCTGACCGGGAGT
CCAGGGCAGGCCGAGGAAAAACACGTAGTTCGCCGCAGTGAGGTTGTTGCCCGTCCCTGCGGCCGCCGTAGTGCAGATGA
ACACTCGGCAGTCGGGATCCTGCTGGAAGCGATCTATCGCCTTCTGCCGCTTGGTGAGCGAGTCATTGCCCACCAGCGTG
ACGCAGCCGTGTCCGGCCTGCTCGCAGAGTTCCTTCAGCGCAGCCACGGTCGGCTTGAACTCGCAGAACAGGATCACCTT

GTCCTCTGCGTCGAGCTCGCTCAACAGGTCCATCGCCACGCGAACTTTCACCGTTTCCAGGTAACGCCGCAGCGCGCCGA
GTCGCGCGAAGACCGGTCGGTCCTCGAGGCGCAGCACGTCGTATTGCTGGCGTTCCTCGGTGGAGAGGGCCACCTTCAGC
AACTGCCGCTGCTTGCCCTTGAGGCTGGGCAGCACATCTTTGCGCCTGCGCAGCATCCAGTCACCCAGCTCCGCCCGCAG
ACTCTGGCGGAACTCCGGGTTGCCGGCGAAACGGTCGCAGAACTCTTTCAGCGGCAGTTGGCCGATGGGGTGGCCTGAGA
GGCGCAGCAGGGTGTGCAGCTCTGTCTCGCGGTTGAGCACCGGCGTGCCGGTAAGCAGGTAGCGGTTCGGCACTTGGGCG
GCAATGTCGAAACCGTGCCGCGTGCATTGCGCCGTCGGTTCCTTCATCCGCTGCGCCTCGTCGATGACCATCACGGCGAA
GCGCGAAGCGTTGGCGACGAAGGGGCTCAACTGCTCGTAGTTGACGAGGATCCACTGCGCCTCTGGGGTGTCCTGCTGGA
TGGCCACGGTGGCCGAGGGATAGACCTCCTGGATCTCCCGCTGCCAATTGATCAGCAGGGTAGCCAGGGTGATGACCAGG
ATTGGTCTGCCCGCCGCGCGGATCGAAGCGGCGATGACCGCCTGGCGGGTCTTGCCCAATCCCATGTCGTCGGCCAATAA
GGCGCTGGTTCTCTGCAGCAGATGAGCGATGCCGTCAGGCTGGTGGGCCAGTAGAGAGTAGCCCTGAAGCGCCGCCTCGA
TATCAGCCGAGCTGTACTCGGTGCGCTCGATCTCCGGCACCGCGGCGAGGTAGATGTCAGAGGCTGACTCGTCCTCCAGG
GATGGCGCTGCCGGTTCCTGCTGTGGACCGCCGATGCTCATGCTGGGCAGCTCGGTCGACGGCGCGAAGCTGCCGTCGCT
GAGCAGCTCCTGGACAGTATCCAGAATCTCGAATTGTTCCTCGGCGAGGCCAAGCTCCAGAATCAGGTTGCTTCGGAGCA
GCTCGGGAGTGGTGTCGATGCGCCAGGAACTCGATGGGCGCAGGAATACCCCGCGCATCCGCCGAGCCACCGCCACAACA
CCAGGATGGAACTCGCCTGACAGTAGGACGCCGCCCTCGGCCAGGGGCGCGATGCGCAGCAGCATCCCCCAGGTGAAGGC
CTGTCGGTCAAGCGAGGCCTGCGCGGATTCGACCAAGGAGCTGAACTCCCGCCAGGACTGGGAGTGAAGGCCTCCAGCTA
GCTCGGCGAGACGGTTGTACAGAGTGTCCAGTTCAGGCAGCAGCTTGCCTTTGGGGACCCGCCAGTAGCGGTGCAACGGA
TGCTCCCGGGCGCTGATGAAGAAGCCGGACAGCTCCAGCAGCGCGCGATTGCTGCCTGCGATGTAGGGGATCTTCAAGCC
GAAGTCGTAGCCGTCGAAGATAGCCCGCTGCAGAAGCGTCGAAGGTGTCGCGGAGGTACGATTAACCAGGCGTGGACGGT
TCATCGGTCTTCCAACTGAGTGAGCAGCCACTTCGCCGTGGCTTCGGCGAGGCCCGCCTTCCAGCGACCATTGGCGAAGC
GCGTCATGTGGACGGTCATGTCGGGCCAGTTCAGGTTCGTCGCTATGGCTCCAGAGCGAATTCGTTCCATGAACCAGCGG
TATGCCTGGCCGGCAGGCGAGTCGGTCGGCACGGGGGCATTATCGAATTCGCGCAGTTGCACCGTGGAAAGGTGCTTGGG
ACTCCAGGGCCTACCTATAGGCTCGCCATGCTCACGCCTGAGCTCCAATTCGGCAGAGGCCAGCGGACGACCAGCCTCAC
GTACTGAGAGCAGTCGGCAGTCTCCTCGGTAGGCGGCCTGGTCGTAGCTATCGATGCAATGTGCCAGAGCGATATGTTCG
GCATATAGGTCGGCAGGACAGCGGAGCTCGACGATCTGCAGTCCATTAGGGCAATCGATAAGACCGACAGACAGCATTCC
CTCCCAGGTGGTGTACTCGTCGGTGTCGACCAGATCGGGCAGGCTGCTCTGCACTGCGGCACGCAGCTCCCTCTGGACCT
GGTGGTAGTCGTCCACCAGGTTGCCAGCTTGTCGGTAGGTACAATGACCCAGATACGCTTTCAGGGCGCTGCGCGCCTGC
CGAACAACCTGTGAGTTCCCCTGGTCGAGGGCGGTATAGAACTCCTTGATGTCTCTCAGCCGTGCAGTGATTGCAAGCCA
TGCCGGATCCTGCCAATCCGACGGGCAGCCGTTGAAGAGACGGTTGTAGTCGGGCTTGGCGTTGTGAACTTGCCAAGGGA
TCGCATTGTAGGCGGCATAGAATGCTGTCCATTGAGCGCGAGTGATCGGCCGCCGGTTACCTAGCGATGCCGCCAGTAGG
AGAGCATGCCATCCCGAATGTCGGCCTTCCCGTTTGCGATGGAAGAGAGCTCCTCCCGCACGGCCGGGACTGACGTGACT
AAGAAATCGAATCGAGGCCCGCGGCGCCTGAAATAGCCAGGCGAAAAAGTTGATGACCGAAATTCCTTCGTCCGCGACCT
GCCCAAGAATACGGCCCATATCGAGGTACAAGTCGCATGGCATGGCGGCGGCCTGCGGTCGTTCACTGTCAAGCTGGGGA
AAGGGGCAAGGAAGGTAATGGCGGATGCTCTCGCCGTTGTCGTTGGTCGTCGTGTGAGGCCAGACTCCCTCCTCGCAATC
CACCAGTAGCGGCGTCAAGACCGGCTGAGCCTTCAGCATTTGCAGCCGGCGCACCTGGTCGCCTTGAGCCAGCCAGTTGT
AGAGCTGGGGCGAGAGGCACCGAACCGAGCGGATGGCGAACAGCAGGTCTTGGTCCAGGACCTGTCTGAAGGTTCGCAAG
ACCGCTCCGATCTGCTTCAAGTAGAACTGGTGCCTTGGCGCGTAGTAGCGCTTGTTGCTGTGGCTCGGATACACGAGGCG
ATGTTGAAGTCCGAGATCGTCCCTGGCCATCTGTGTGGCTAGGTCGTGCGCTCGCTTTGTGTATGAGCTGAGTACAGCAA
CTGCGTCGTCGAACTTTTGAAGGCGCTGTAGCAGCTTCAAGGCCCCTGCTCGAGACGGCGTGCTTGCCGGGATACGATGG
GCAGGCTTGCACTCCCACGCACCAATTAGTTGACCGTCGAATACCTGGAGTAGCGGCCCTGCACGCATGATGGTGTAACG
CACACCGTTGTCAGCAGCGACTGTGATGCACCAGCACCTGGTCGTAGATGTGCTGCGCAGGTAGTGACGGATGAACTTGG
TTCGAGCTGTCCCTGCGAGGGAGAAATGTGTTGCCCAGGTGTTGGCGACGGCGAGCTGCCGCTCGCTGTAGCGGGCCAGG
GATGGATCCTGTAGCGGCCCGCTAGGGGGGGCGTTGTCGAGAGGTGCCACGCGCATGCTCCTGTGTAAGCAAGCAGCACT
TTGGCCGGTCGGTTTACGCTGAATCCATGTGAAATGGTCATCTCAAGCGTGTTGGATTTCAAGGTTCTACAAGCTAGCTG
ACTGAAGGTCTTATTGGGATGGATCAGTAAATTCTCTTGTAGTCAGCCTCAAAGTCTCCTAGTGCAGCAAACTCGTCAAG
GATAGCTTTACTTAAAAGAATAAGATGTATTGTCCTCGCTGGGTTGTCCCAGTTGGCACCACCATGCTTGCCTCCGTGAA
ATAGATTGTTTCGCACGGTCTTAAGTAAGCGAGCGACTCTAGCTAAGTCGCTTTTGTCCTCATCTAATTTAACCGGACGC
CAACCCAGCTCTCTACCGGGCAGGACTATTTGTTGCTCTGGACTCTGCTCGATTAGTGCTGTGGCTGATTGGGAAAGAGA
GTATTTGTCAGAATGGTTTTGTACAAAATTTTCCCATCCCGGCTCTGCCCTAGCTCCAGGTTTGTAATTTTTTAAGTAGC
CATTCTCTTTGAGGCTGAACTCAAATCGCGAGAACCAATAGAAAAAATCAAAAGCTCGCTCTGTGATTTCGGCAGATATA
TTCTCTTTTCGCATAAACCCTCCTTGATAGCTAGCAGCGAATTGGGCCTGTACCATTTTGCCTGCACTGCGTGACTAACT
ATAGTGGCTATAGCCCGCGCGGGTTGTCGAAGTCAGGCCGGCCGCTCTTGGTGGGGTAGCTAACCTTGCACCCTGTTGTT
CGATAGCCGCTGCAACCCCAGAAGTCGAAGCCGCCTTTGCCGGCTTTCTTGCGGTGAACCAGGGGCAGACCGCACTTGCC
GCAAAGGAAGTCGCTGGTCGGCGCCGAGGAGATACTGGCGGTGGAGCTTGCGCGTGGTTCGTCCTGTGCGCTGGGGAGCG
CGCGGAGTGTCTGCAGCTCTGTCTGTAGTAGCTGGTAGAACTGCTGGACCACGTTCATGTAGGTGTCCTCGCTGTTGGCG
ATCCGGTCCAAGGCGACCTCGAGGTTGCGGGTGAAGTCGAGGCTGAGGAAGCTGAAGTTGTGCTCCAGGAGCGCGATGGT
GGCCTCTCCCAGCTCGCCGGGGATCAGGCTTCGCCCCTTCATCTGCACCTGGCCCTTGTCGATGATGTTCTTCACGATCG
AGGCGTAGGAGGATGGCCGACCAATCCCGCGGCGCTTCATCTCGCCCACCAGGCTGGCGTCGGTGAATCGCTTGGGAGGG
GTGGTTTTCTTCTCCAGGACCTCGCCGCTGTAGACCTTGAGTATCTGGCGTGGCTCCAGCGCCGGGATCGGGATGGGGTT
TGAAGGCGTTTCGTTCTTCTGCTCCTCGGCGTCATCACCCTGCAGCAGTTTTCTCCAGCCAGGCACGTTCAACAGCTTCC
CTTTGGCGCCGAAGCGCAGCGGCTTTTTGTCGGGGCCGACGCCCAGGAGGGTGATGGTTCTCACTGCGTACACGGCAGCT
TCGATCTGACTGGCAAGCGCGCGGACTCGAATGAGCTGGTACAGCGCCTGCTCATCAGCAGTTTCACCGGCGGCAGCGGC
CATCCAGTCGGTAGGGGTGATGGCGGGGTGGCCTTCCTGGGCGTCTTGGTCCGCTTTGAACATCCGCTGTTGCTCAACAC
ACTTCAGCCCCAAGGCTTTGGCGACAGCACGGATATCCGGCATCGAGTCCTTCGAGATATTGGGGTTGTCCGTCCGGTGG
TAGGTGATGAGCCCCTGTTCATACAGCCGCTGGGCGACCTTCATCGTCTTGTCGGGTGACCACTTCAGCGCATTCCCGGC
GGCCATCTGGAGCGAGGAGGAGATGAACGGTGCCGGAGGCGCATCGGTTTCTTCGCTATCAATGCAGGTCTCGACGATGA
CATTACGTATAGCCCCCACCAGTTCTGCGAGTTGACGATCCTGGACATACGGGAACTCCTCGCTGGCAAACACGGGCACT
GGATCCCATTCCGCCGTCCAGGTACGGCCGTCGCTGGGCGAAACGAAGGTCAGTTCCACCCCGAAGTGCTTGATTGCTGT
GAAGGCGCGGATCTCCCGCTCTCGCAGGACCACCAGGTACACCGCGACGGACTGCACGCGCCCGGCGGTGGTCGGCCTAC
CCATCACGCGCCGCAACTCTGGCGTGACCAGATACCCCACCAGGCGATCGATGACGCGACGGCATTCCTGCGAGGCGACC
TTCGGGAGGTCCAGGCGACGCGGCGAGCTGAGTTCGGCGGTGATGCATGACTTTGTGATTTCTTTGAAGGCAACGCGCTT

GTAGTTCTTGATCCCGGCAGCTTGGAGGATGTGCCAGCCAATGCTTTCGCCTTCGCGATCCGGGTCCGATGCGACGTAGA
TTTCCACGGCTTTCTGCCGCAGCTCCTTCAGCCGTGCGACGGTTTTTTCCTTGCCCGAGAGGATCTGGTAGTGCGGTTTG
AAATCCTGGCCCACGCCGACGGTGAGCATCTGCGGATCCTGCCCGTGAACGGGTAGGTCTCTGATGTGGCCGGCTGTCGC
CTCGACCTGCCAGGTCACGTCGGGACGAATCGAGGGGAGAAGGGACCTCAGCTTTTTCAGCTTGCCCGGGGCTTCAATGA
TTACGATGTGCATCGGCGGAATCTCTCGGAACAGTCGGTCGAACTGTCGCGATCAGAGAGCGGGCCAACAATGCGAAATA
GACGGTGCGTTGTTGCAGTTTTCGGGGCAACTTTCCGCGCATCGGCGAAAAAATATCCGTTGCCGATCAAATCGGACACA
ACTTACCGATTCCTCAAATGCGGTAATCTTGTCCGTGTGGCTGCTGCAACCATCTACGGTAATCGTGGCAACTGGGTCGT
CGCTCGCATAGATTCAGACAAGGAAATGATCGTTCCTGGAAGGACATCAATCGACCCCAATCATCCCGATTGGTCGTAGG
AAGGAAAATGGTATGGATCAAAGCCTTTGCACATGCATGCCAACGCCAATCGTCAACCCCAAGGAGCTGCGACTGTGCCA
CATGTTAGTCGGTAGAACTTTCCCGATAACATTGATCGCAGGCGACCATTGGTTGAGCTATGACGGCAGCGCCTGGTGGG
TCGATGCGGATGAGCCCGCGACGGAGGACGAGGTGGCGGCTCTGTTGGTCAAGGCTGGTGGTGTCACTACGTGCTGGTGC
GGATAGGAAAGCGCATCAGGTACACCCAAGCCCCTCCTAACTCGTGGGGCCAAGGCCGTAGGTGAGCCACGACTTACTTT
GCCCACAATGCGAAATGGACGGTGCGCTGTTGCAGTCTTCGGGAGGAGCTGAAGGGTAAAGTCATCAGAAGTCGAGCTGC
TGTTGCTGCTCCTGGGCAAACCCTTTCAACCAGGACGCCCGCATCGCAGGATGACTGTAGGGGCAGTCGAACACCCGTAG
CCCTTTCCGCGCGGCCTCAGCACCAATAGCCCGTACAGTCCGGTGGGGATCGTTGAAGTGTCGTCTGTCCAGCTTGTCGG
TATCGGTCGCTCTCTTTGACACACTGCACCTCGTACTGGTCGAGATTTCCGATTCGGAAGCCCTTGCCACCAGAGACTTC
ACGAACGGGCGGCCAAAATTAAGGATGCTGGTCTATCAGCTTTTGCCGCCATAACTGTTCATTGCGGTTGCTCCAGTCGC
TGGAATGCTTGTCGAGGCTGCGGGCGGCGAAGAGAGGTTCTTGCCACCTGCAACTTCACGAGTGGGCGCGAAAAATTAAG
ACCGGCGTTTCACCGTCGCGTCGGCAATGCCTTGGGCCATCCTGATGAAGGATGCGCGATCTTCTTCCTCCATAAGCGCC
ATGCACTCTAGAACTGATATCTCGTCGGCGCTGAGGAGCTGCCTGGTCAACGGCTCATCGCGATCCCGAAATGAGCAATG
TGAGCCTGCTTGGTCATCTTTGCTGTTCATGGCGGAGTTTCCTGCGGGGGGCTTTGTCGAAGCTGCGGGTGGCAAAAAGT
GGTACTTGCCACCCGCAGCTTCACGAATGGGCGGCCAAAAATTAAGCCGCCTGCTACATGGGGATGTCGTCGTCGAAGGC
GCTGTCGTAGTCCTGCTGCGATTGCGCATCTTGCTGAGCAGGTTGCTGCGGGACGTTCCGAGACTGTTGATGCTGGGTTG
GCAACAGGGTGACCTCGGCCAGGCGATGCGGAAGGATGGAAATGCGCGAGGCTTCGACCTTCAGCGCCTGGACTTCTTCG
CCTGACTCTTTGTCCGGCCAGCGGTCCATAATGGCCCTGCCTTCGACCTTGACGCGCATACCTTTCGTGAACAGTTCGGC
GAAGCGCTGGGCATCCTGATGCCACCATTCGACGCTGCACCAGAAGCCGCCGCGATCCTTGTAGCCACCTTGGCCATCGG
GAATCGAGTTGTCGAACATCACGTTCAGTCGCAGCAACTGCCGCGGGGGATTGTTGCCGTTGGGGAAGCTGCGGTGCTCC
GGCGCCGAGCCAATGTTGGCTTCCCAGAAGACGGGGGTAGCCATGATGAATCTCCAGTCAGTGGTGGTTGGTGCTCTGCT
GCAGTCGGCCCATGAACTCGGCCTCGGCCTGGCCCATCTTTTCCGCGCACTCGGCGGCCTGCTTGCCGATGGAGTGGATC
AGGCTGATCTGCATGTTGAGGGTGATGCGCTGCAGCTCCATCTCGTAGAGGTCTTGCAGCAGGTATTCCGGGGTCTTGTC
CTGGCGCAGCATTTCCGTCCACACCCCGACGCCCATCCGCCGGGATGCAATGTGCTGCCAGCGGAGAAAGGTTGTGCGGG
CGGATGTGCGCTGCTCGATGAGTCGGGTGGGCAGTAGAGTGAAGGGTGGGCGACGAACCTGGGCCAGCAGTCCCTGGGCA
AGTTCAAGCAATCCCTGCTCCATGGCTTCGCACTGTTCCGCCAACTGGGCCAGCTCCCTATTACCTTTAAAAGGCTTTAA
AAGGCCTTTTAGTGAGGCACTGTGTTGCAGCGCCTGGAAGGCAGTTTGTTGCAGTGGGCGGAAGGCACTTTGTTCCAGAG
CGCCTTGCTGGTCCTTATCGAGCTGCCGCATTCCGATCAGTCCTCGAGTTCGACACGGTCAGCGTCATCATCGGCATCAC
CGTCAGAGCGGCCCCGGGTGATCGGCGGAGCGAAGTTCGAGCGTCGAGTGCCCTCCAGGATGTCCTGCGGGATCTCGCCG
AACTTCTCGTACATCTTCCGCGCGGCTTCGGCGCGAGCATTGTTTGCGGCGAAGTCGTCGCGAGTGGCGCCCGAGAACTG
GTAGCTCTGGGCCAGACCAAACAGGCTTCGCAGCACCGACGCACCTTCGTCGAGCCAGACCTCCATGTCGCGGCGACCGA
CCAGCCCGACGTGCTGGGCCAGCAGGATTCGACGGGCGAGTTCGTCGTAGTTGGTCAGGAGATAGACTGCCTTGAAGCCG
AGAGGGTTGGAGATGAACAACGGGACCTTGACCGGTGTAACGGACAGGTTCTCGGAGATATCGAGGGTCGCGGGCAGCTT
GGCCATGACGTCATCCAGTCGCTGGTCGATGGCCTCGAGTGCGGCTTGGCAGCTCTCGACTTTCTCTTCGATCTGAACCA
GCCACCAGTCGGAGTACGGATCGTCCTGTGCTGCCCCGCGGTGCATGCGATTCACGTATGCGCAGAAGCCGGAGAGGCCG
AGGATGCTGTGCTTGGCATCCGACTTCTGCCGGCCGGTCCAGATGCGGGCGGCGTGATAGGTGTGAAGCGTCAGAGCGAC
ATCGCTGCGGAGAGAGCCGATCTGAAGCCGATGGGTTTCAGCCATGGTGCTGTGTTCCTTTATAGGGAAGGGGGTCGGCC
ACATCGTGCTGATCGGGGTAGGAGCGCGCAGTAGGAAATACATGGTTCGCCGCTGGGGAATTTCCCATGGCGACTCGGAG
CAAGGCACCCAGTTCGCCGCCGGCGTCGAAGTCAGTTCCTAATGGAGTGGTTCCAGTGGAACCACCCAACTCGCCGTCGG
CGCCGAAGTCAGTTCCTGGTGGAGTGGTTCCAGTGGAAACACTCATCTTGCCGTTGGCGCCGACTTGGGTATCTCTACCG
TTGGTAATCTGGTTGGCAGTTCATGTAGTTGGCTTTGTTGCATAAAGCTCCCCACATCGCTTTTTCCGCCATTTCAAAGC
GATACATCTCAACCGTAATGGTCAACACGGTTACTGAATACCCGAACACGATCAGCATTAACAACAGCCTGAGTTCTCGG
TCGTTCCACGCATCTGTAAAAAGTGCAGAAAGAACGAGAACGACCAAAAAGAAAGGGGCCATGCTCAAGAACATTGTTTG
AGTTCCTCGTGAGTGGTGATGCCTCGCTTTGGCGAAGTGTTTCGGCCCGAAACACCCATCTTGCCGCCGATACCGAAGTG
GATCCCTGGCAGAGTGATTCCAGTGGAATCACCCAGCTCGCCGGCGGCGCGGAGGTCGGCTCCTGGTGGAGTGATTCCAG
TGGAATCACCCAGCTCGCCGGCGGCGCGGAGGTCAGGACCTGGGAGAGTGATGCGGGCCGCATCACCAAAAAACACGACG
CCCGACCGAGACTTTCTCATGAATCATCACCCCGCTCGGATGGCACGCTCGATCCGTGCCTGGGCCGAATGAGGTCCTTG
AGCCGGCTGAGGACCTCGCTCCCTGTCCGGGTATCCCCCCGGGGCTGGACCTGAGGCTGTGTAGGCGCTATGGGGCTTGA
TGGTGCCGGAGCACGAATGGGGCGTTCCGTTGCGGGGATGGTTCGCTCAGCCGGTTCTTCCGGAGCCCAAGAAGCGTTGA
ACTTGCCCTGGACTGCACGCTGGATGAGCGTCATGAGGTAGCCGAACGGATTGCTGACTGTTCCGCCGGCACAACGGTGC
TGCCACTGCTCGATGATGGGTAGCCGAAGATCCTCGGACACCCGTCTGAGCGCGCTTACTGCCTGGATCCGTTGCTCGGC
CTCCAGTGCGTGCAGGAGATCCTGCCAGTTCGGATGGGCTTCCTCGCGTGCTCGCGGTTGTACTGGCTTTTTACATACAG
AATCTTGTTTGTATGTATACGTACTATATGAGTTCGGCATCCGAACCAAGGGCGTACTCGGCGGATTCTGTCGGGGTTCG
GTATCCGAACTCAGCGGTAGAGCGCATTGCTTGCCGCTGATTTCACTGAGTTCGGCATCCGAACTCGGGGTGGAATGAGG
TAAATCCGACAGAGTTCGGATGCCGAACTCTGCCGCTGGATCCGCTTCGAGCTGACTATCGATAGCCCATTGCTGGTGGT
TCAAGCGTCCCTCGAGCTTCTCCAGATGGGTAGGGAGGCGGCGACCCACGTCCGGATCATCCCTGAACTCCCGCCAGATG
ATCTGCCCGATTTCGCGTATGGCTCGGTTGCCGTGACCGGTGGATTGGCTCAGCAACTGCATGTAGTCGGTGTCCAGCTC
CAAGGCTTCGGCTGGAGAGACAGGCTCGTCGTGAAGGATGTAAACGTTGCCCTGGACCTGTCCGTTGAGGTCGTTGCGCA
GGCGTCGGCCGAGGCTGAGCCAGCGGGTCAGACGAAGGACCGTGAGTGCCTTGGCGATGGTTTCGCGCGAGGCGATCTTG
CCCGGCTGCATACCGAGATAGGGGCGCAGTTGCTCGTATGTCGGGAACGCGGTCAGTCCGTCGTCGTTGATGAGCAAGCG
GAAAACTTGCCAGCAGTTCCGCTCCAGCGGAGTAAGACGCTTATCCAGGAGGAGGCGCCGTGGAACGGTTTCATGAGGAT
TGCCGCTGAAAATGATGCCGGAATATTGTTGAGCCAGCGGTTCGCCGCTTTCGGCAGCAAGCCGCTTCTTATACAAATGG
GCCTCGAGATGCCCCGAGGCCGAGTCCAGCAGTGTCGAAATGGGAAAGCGCGACAGCCTCATGGTGCTGTGTCGCCTTCC
GACTGAACCAGGTTCAGTAGTTGTTGACGTTCTGTCTCAGCGTTCACCTGCGTTGTTGCGGATGGGGCAGAGTGGGGCAG

GTGTGAGCTGACGCTAGCGAGGTAAAGAGTGGATTGGGACTTTTGCAGGCCCGCCTGGCTCTGGCTCGATTTGCCAGACG
GATAGAGCCCGTCTTTCAACCAGCTCTGGATGAGAGACCAAACAATGGCGAGGCTGATGAATTCTGGACTGTCTTGTTCG
ATTCCGGCGTTGATTTCCTCTGTGAGGATCATCGCGATGTCCAGCAACTCCATGCTGTCTTCAAGGCGGACCTGATTTTC
CTGCATCAGGTGCTCCCAGCGGAACCATATCTGGCGCTCGATCTGTTCATCGAGCTGCCGCCAGCGACCGCGGCGCGGGT
GGATCTCCAGCATCAAGCGCTTGGTGGCGGTGTCCTCCGGCGACAGACCGAAAAAGCTCTGGATCATCGTTGTGGTGGCT
CCCAGCTTCAGTGCGCGTTCGATCTGCAGGTCTTCCTGCGCGCTGCGCTCGGCTGTGGCCAGGATTTTCTCCATGACGTC
GATGTTCACTCTGACATCTACCCAGGAAACCGGGGTATTGGTCAGGATGCTCATGACGGCAGGCTGTTGGAGCTGAGCCA
GAATCTCCGGCCGCAGCCCCATCTCGGCACAACGCCTGAGCTGGCCGTTGCGCAGGTGGTTGAGCACCTGCGTCAGGACC
GCTTCATTGATCGTCGACTTCGACATGTCGATCCCTCCCCAAATCGTGCGAAATCACTGGTCAGTAGCTGCTTGCTCGGA
GGCGGCCGGCGGCAGCTCAAGGTCAACCAGGCGTCGGGCAAGACGAATGATCCGGAACAGTTTCACCAGGGCTTCGTCAT
TGAGACGGCCTGCCTGGCAGGATGGCTTGTCACTTGTGGTGGGTTGGCCCAGCAGCAACTGAGCGAGGCCAGCTGAGAAC
TGGAATTCACCCAGTGCATCGCGGCAGGTCTGTTGATCGTGAGTGTTCAGAACGACTTGGACCTGCCCGCTGACCGCTTG
AAGCAGCGTCATGAGGTGCAGCGCTGAATCAGTAATCTCATGGTCTTCATTGGGCTCGCGATACGTGTAACCCACACCGC
CCTGAATCCTGACGATCTCGCCTGGGGCGCCGACATGCAGAGCGATCTCTTCAGCCAGATCAGCGATGTGCTGTCGCAGG
ATCTCGGGGGTATCGATCGACCGCTCGATGTACCAGAGATCGGTGATCGGGTGCAGGCCACCGACTTGCACTGGGATTGC
CTGCAGAGCTGTTTCATCGAAGACAGGTAGATGTTCCCCCTCGAGGCCGGCCAGCCGTTGGCGAGTCTGCTGGATCTTAG
TCGATACCGGGCTCACGATATGGCCTGCCAAGACCGCCGCACGCTCCTCATCGGGCAATTGCTTTTGTTGGACAGGTGGT
GGCGGAGCCGGCGGGGGGAGGGGATTACCCGGGTTGCTCCTTGTCCTGGACGTTTGGGGATCTGCGGCGGGGCTCTGTGT
TTGTGCTTGTCCAGTAGACGACGCAGCAGGGTTTTCCTGCCCAATTGGTGGGAGTTGAGGTGGTGCGGCAGGTGTGGGCA
GGTCGACCAGAGTGCCGCGGCGTTGCTCCTGCTGCCCGTTGGTAATCTCGAGCAGGATTTGGTCATAACGCAGGCCCAGG
GGGCGCTTCATTTGGTCGATGAGTTCGTCCTGGAAACGCTCGAAAATGAATTCGTCAGGGCTACTGTCGAAGATTGCCAG
GGTGTCCTGGAACAGCATTTCGAAGTCAACCCCTTCACCAGCGTATAGACGTGCCCAGCAGGAGGATGCCGACTTTCTGA
GTGACAGGAGTTTCTCGATTTGCGGTTTACCGAGCCCTGAATACAGCATCACAGGAATCGCCGGCGCCAATACCTCAATA
GTGTCTAACATCTTACTGATATGGGATTGAGATACAGGGTAGCCGTCCGCTTTTAGCCGACGTGCCAACTCGCGCTGGGA
AATGCTTTCACCGCCGTTCTCTTGTTCGTAAAGAAATTTCGCCTTCTGCACCCCAACCGCGCGCTCGATGAACTTGAGGT
CGCCCTTCAGATCGTTCTCGGCCAAATGACCGGTCAGCGCGATGATTTCGCCGCGCTGCTTGTCCCAGGGCTTGAACAGG
CAGTCGAAGCTGAAATAGCGCTCGTCTCCGGTCTCTTTGTAGAGGTCGTTGAGAATTTCCAGGCGAGTGTTTCCGCCGTT
GCGAATGCGGTATTTGTCCTCTCCAGGGCGTCGAGTCACTGGTGGCGGCGTATCGAGGCCTCGATGTCGGATCGATTCTT
TCAGCTCATCGTATTTCGGGTTTCGAGTGGTCCGGGGGTTATCGTGCCAGGGGAGGACGTCGTGGAGTGTCAGCCTCATG
GGCGTGTCGCTGATCGGGTCCGACAACTGCTCCAGAACCTTGGGCGGTGCGGTAAAACCAGGGGCAAGTAGCCGGTCTTT
CAGATCTTGCGGGTTGATCTTGGCCATACTCATGCCTCCGCCACCTGAGGCGTGTTCGCCGCCGCCAGCTCAGCCTGCAG
GCGCTGGTTCTTGGCGAAGAGTTCGTCGACCATTCGGTTTGCTCGCGCGAGTTCTTCTTCGAGCGCGTTGCTGCGCTCCA
GTCGGTAGACGATCTCGCTGTTGGCCGAACGGTGGTTATCCGCGGCTTTTTGCTTTAGCTGATCGCGCAAGCCCAAGGGC
ATACGGACAACGAACTTATCTTCGGTTTCTCGGGCCATGGTCATTCTCCAATTTGGGCCTTGCGCCGCTTGAGCTGGTCG
CGGAGTTGAGTGGGGAGGCTGCTGATCTGCAGCGTGTCGCGCTCCTCGTCGTACTCGATCCTGGAGCCCAGCAGGTGCGC
CTCGAAGCTGATCGACAGGCCTTCGGCGCGGCCGGTGAAGCGGCGGAACTGGTTGAGGGTGCGCTTGTCCGCCGGGATTT
CCGGCGACAGGCCGTAGTCCTTGTTACGGATGTAGTCGTAGAACGCCCGCGGTTGCTGGTCGTCCATCAGTTCCGAAAGC
GCGTCGAGGGTCATCGGCTCGCCGATGCGCGCCTGCGAGGTGGCGTAGTCGACCAGCGTCTCGGTCTTCTCGCGGGCCTG
TTCCTCGGCCATGTCCTCGCTTTCCACGAAATCGCTGAAGGCTTTCAGCAGGGTGCGCGTCTCGCTCGGCGAATCCACCC
CTTCCTGGCAGCCGATGAAGTCGCGGAAATAGTCGGAGACCTTCCTCCCGCCCTTGCCCTTGATGAACGAGATGTACTGC
TTCGACTGCTTGTTGTTGCGCCATTCGGAAATGTTGATCCGCGCGGCCAGGTGCAACTGGCCGAGGTCCAGGTGGCGCGA
CGGGGTGACCTCCAGCGACTCGTTCACCGCCACGCCTTCGCTGTGGTGCAGCAGGGCGATCGCCAGGTAGTCGGTCATGC
CTTGCTGGTAGTGGGCGAACAGGACGTGGCCGCCGGTGAAGAGATTGGACTCCTCCATCAGCTTTTGCAGGTGCTCGACC
GCTTCGCGGCTGAAGCCGACGAAGTCGCGGTCGCCCTCCAGGTACTCGCCGAGCCAGCCGCTGAACGGGTAGGCCCCGGA
CTCGCCCTGGAAGAAGCCCCAGGCCTTGTTCTTGGCGTTGTAGCTTTCGTTGAGGTCGGCCAGCAGGTTCTCGATGGCCT
GGGAGTCGCCCAGCTCGGCGTCGCGCGCGTGGAGCACGGCCGGGGTGCCGTCGGGCTTCTTCTCGATCAGGTGGACGATG
GCGTGTTTGATAGGCATCAGAAAGACTCCTGGTTCAGGGAACGCTGCCTGTCGGCTCGACGGAGCAGGGCGCGGGCATTG
AGCTCGCCGCGTGCGTCGCTTGGCGTGGAAGAAATGAAGATCGAGGCGCAGCCGATCTTGCTGAGCTTCAAGTGGCCGCC
TGCGGAGCGGTCGACGCTCCAGCCTTCTGCTTGAGCGAAGGCCACCAGCCGGCGAACGCTTTCCCGGGCGCCGCGAAGGA
AACTACGCATTCGTGGTCTCCTCGCTGTTGAGCTGCAGCCAGGTCTCCATCTCTGCAAGGGAGCGGGGAAGGCTATCGGT
GGCGGCGGCATCCAGGAGACGTTTCACCAGCGCCTGAAGGCTGCGACGATCGCTCTGCCGGCTGCAGGCCAGGTCGTCAT
GCGCGGAACGCAGCATGCGGACCTCTGTGCGCAGTTGAGTGCATACGTCGTCAGTTGCTACTGGTGCCGTTGACGCGACA
GGTGACCGCGACCGTGTCCTGAGGCGGTTTTCCCACTCTTCCCTTGCTGCCGAGAAGCGATGGCCACGACCTACACTCCA
ACCGCTGGTCCAGTAGCGCGACTTTCCGCAGTCAGGGCAGGGCTGTGCGCCGCTGCCATCCAGGTGATAGCCGCAGTTCA
GACAGATCCACGAAGGACCTTCACGCAACTTGAGCTCAAGCATCAGCTCGACTTTCTGGAAGTGCTCGAGCGGCCGGTAG
GCTGTGACGTCAGGCATCGCGCGCTACCTCCTTGGGATTTCTACAGGACAGCGCCCAGCGGCTCTCGATGCGCTTTTCGT
GGTCGCCTGCCTGTACCGGCATACGGGTGATGTATGCGTTGTCCTCGCCAAACAGTCGTTCGGCAAGGCGGGAGGCTGCC
ACCTTGTCGCTCCAGGTGCAGGTCGCACGCTGCTTCTGAATGGTATTGGTCCAGAAACCGGCGCCATCGAACTTGACGAT
GCACTGATAGGTCTGGGACGCCTTCATGACTCGGCCTCCTTGCTGAGTGCATCACCGACCTGAACGGCCACCAGTTCGGG
CAACTCCGGCTCATGCCCGGGATAGTTGTTGGCCAACTGATGCAGGAACTCGCGCGCGGCGTCCTTGCTGCCTGCAATGC
GGAACACGGCGGTGAGGTGCTGGGAGTCCTGAGTGCCGCAGCAGAGGTGCGGCACCAGGCTCACTTTCAGCGAGCAGCTG
AACAGGTGTCGCGGGATCCATACCGACCAGAAGTAAGCCGAAAGCAATTGCCCCGGAGCGCGCGTAGTGACCACCAGAGT
CCCGGCCTGGGTGGTCGAGATTTCGATGTACCCCAGTTCCTCCAGCTCACGAAGCTCTCGCTCCATCGAACTGGGTGCGA
CCGTGCTCTCGTCCCGCGAAAGGCGCGAGATCACTTGGCTGAGGGCGTCCAGCTCGATGGGGTTGACCCCAACGTGATTG
GCATGAACCACTGCCGAGTCGGTGAACGCGGCCGAGGCGTGGATACTCGGAAAGCGGGCTTCAGTGTTCATCAAAGACAG
CCTCCAACTGATGGACGTCGTAGCTTCCGCAGCAGGGGCAGACCAACACCTCGGCTTCTCCTTCTGGGGTGTTCTCGGCG
GGGAGAACCTGGCGCACCAAATGGTGTGGGTGGTGCTGGTCACAATTCATGCACTGGTGGGTGACCAGGGGCTGGGAAGT
GTTCATGCATGGGCTCCTTGCCAGGTCCCCGGCGCCAGGTCGGCAAAGCGGGTCTGGTTGGCGATGAAGGCGGTGTGGAC
GGTGCCGATCGGACCGTTGCGATACTTGCCGATGATGATTTCGGCGATGCCCTTGGCCTCGGTGTTGGGGTGGTAGACCT
CGTCGCGGTACACGAACATGATCACGTCCGCGTCCTGCTCGATTGCGCCGCTCTCTCGTAGGTCCGCGCAGTTCGGTCGC
TTGTTCGTCCGGTTCTCTAGGCCGCGATTTAGCTGGGATAGAGCTACGACGGGACAGTCCATCTCCTTGGCCAGCGCTTT

FIGURE 2O

CAGCGAGCGGGAGATTTCCGAGATCTCGGTGGCTCGATTCTCCAGGCCTGGGCAACTCATCAGTTGCAGGTAGTCGACCA
ATATCAGCGCGGGGTGTCCGTACTTCCTGGCCGCCCGGCGAACCTTGGCGCGCAGCTCTGTCGGCGTGAGGTTGCCCTGA
TCGTTGATGACCAGCCGGCTGCCATAGTCGTTTATGCGCTGGATCGCGCCAGACAGTCTGGGCCAATCCTCTTCTTGCAG
TTGGCCCTTCATCAGCTTGCCCAGGTCCAGGTGGCCGAACAGGGCGGCAAGCCTGAACAGCAACTGCTCTGCCGGCATCT
CCATGCTGTACACCTGAACAGACTTCTGTTGGTCGCTCTGGAGCGCGGTGTCGACCAGGTTGAGGGCAAACGACGTTTTG
CCCATCGCGGGGCGGGCACCGACGATGATGAGATCCGACTTCTGTAGTCCGCCGGTGAGTGCGTCGAGATCCTTCAGGCC
AGTCGGGACCCCCGTTACCGTCACGTTGTTGTTGAAGCGGTAATCGATGGTGTCGACGATCTTCGTGAGACTCTTGTTGA
TATCGACGAAATCGGCGTTGTGGTGGTCCTGGCCAAGGGCCGAACAGCTTCTGCTCAATCTCCTCCTGAACCTCAGAGGG
ATTTGCCTGGTGGTTCGAGGCGGTCCTGGTGCAGTGGTGCCCGAGAGACATCAGCTGCCGCAGGTGTGCCCGATCGCGAA
CGATCTGCGCGTAAGCCTCGATATTGGCCACGGAGGGCGTGTTGTCGGCGAGCTGGCCGAGGTAGGCCAGCCCGCCAGCT
TCTGGAAGGTCTTCGATCGCTTCCGACACAGTCACGACATCAAACGGAGCATCCTTCGCGGCCAACTCGCTGATGGCGGT
GAAGATCAGCCGATGCTCATGCCGGAAGAAGTCCTCCTTCTGCAACTGATCGCCGACAATGTCCCATGCCGCGTTGTCCA
GCATCAGGCCGCCGAGTACCCCTTGCTCCGCCTCGACGGAGTGCGGCGGCGTAAGTTCCAGAACACTCACTTCGAGTCCT
CCCGGCTGATCTGCAGAAACGTGCTTCCTCTGCCGCCCGGGGAGCCGATGATCTCCAGCGCTTTCACTTCCACCTTGGCG
GACTCGATGATCACCGAACCGACATGAGCCACCGCTTTCGCCGCTCGATATCCAGGGGCTTGTCTTCGTCAAGCAACCC
TTCGATGGTGGCAAAGAGGTGATTCCGCAGATCCTCAATCTTGTTCTTCATGCTCGACCTCCTGAATGGTGCGTTTGAGC
TTGCTGATGACGCGGATTGCGCTCTTCAGCTCAGGTGGATAGCGATGGATGGAGTTGCGGCGCATGCGTTCGGCCCGGGT
GATGAGCTCCAGGTTGTCGATGACGACGTTCTGCTTGTTGTTGTCCTTGAAGCAGACGCAATGGCCGGTTGGGATGGGGC
CGAAGTGTTCTTCCCAGAGCAGGATGTGGATGCCCTTCCAGTCCCGGGGGGGATAGCCGGTATCCGAGATCTTTCGTTGC
AGGTAGCCATCAGCGCTGACCCGCGTGCTGCCGACTGGGAGCCATGTGTGGGGCTTTTGCCCCTTCTTGAACTGCGTTTC
TGGTGCGCGTCCTCGCGCGGGCAGCCCCTTCATGCCGCAGTTCCATGGGGTGTTTCCCTTCTTGAACCGAGTTGCGCTGC
CCCAGTCCTGCCGGATCTTCTTGGAGAACTCAGGGCTTTTTTTCAGCCCGAGCCGATACGCTCTGGAGCAGATCTGCTGG
AGCGTTTTGTTCAACCTGGCTGCCAGGACCTCGTTCGGCATATCCGGATACAGGCGCAGCAGCACCAGTTCCTGTTCGGC
GGTCCAGGTTTCGCGGATGCAGCGCTCCTCTCGGTTGCTGAATTTGCGGCGGGCACGCACCAGGGCCTCGTAGGCGATGG
TCGAGAGAGCAGAGCGAGTCATTGGGAACCTCCCTGCTCGAGCGCGCCCTGAACTCGCCGGCTGACACCGCGGTGCAGGC
GAGTGTTTTTGCCGTGCTCCCAGCCTGCTTGCAGCGAGGCCTGGTCATAGGCAAGGGCCTTCGTAAGCGCCGCCGGCTCC
AGGTACTTCAACGCCGGATGGTGTAGTTCGAGGTAGGCCTTGATGGCTGCTTGAGTCGACTCGTCCGGCCTGCCAGCAAA
TTCACGTACCAGCGAACGCACTGCGAGAAGCCAGCCTTCGACGAAGAGCTTGCTGCGACGACGCTTGGTCGACAGCTTGC
AGCGCTTCTGTTGGGCGACATGAGCCGAGCGCGCTGCCACCAGTTGGTGGTGGAGCGTAGAGTAGGCGTGAGCGGCCAGC
TCAGGGCCGATCCCTCGGCCCATGAACTTGAACGTCCAGCCCGCTGGCATCGCGTATGCCGCCAGGTAGTCGCAGCCGAA
GGAACTGGCGCAGATCCCGGCCAGTTCATGCAGCCATGGGGCAGGAGAGCGCCTGGTTCGAGTGGGAACGCACGCGGTTT
CCACCGTGTGGGCCGAGATCTCCGCATCGCTGAGGTCGAACTGCCGCTTCAGGATCGCGGCCTGGCGCAATGCGGTTTCC
GCCTCGTTCGGCGTGGCCCCCTTGGCATTCGCCAGGGCCATCAGCTTCTCGACCTTGGCGAGGACTTTTTGCTTGTCCAT
TACGCGCCCCCCTTGCAGATGGTTGAGCCCTGCTCGGTGGCGGAGCGGTCCAGGCGCTCGATCTCGGCCAGCACTAGGGC
GCCGGCGCGTACCAGGTTCTCGCGCGCCGACTTAGGCTTCCACCACTCTTCGTCCCAGGGCCAGAACAGTGCAGAGCCAT
TCCGAATGCCGCCCGGCCAGTCGGTGCCGATGCCGCCGGCATGGAGCGCGTAGCAGCCGGCCGCCAGGGCGATCAGGCCA
GCGCTAGCGTCGTCGTTGCTGGTATCGAAACCTTCGACGGTGATCTGGCGCCGGCGCTCGGCCTGTACATCGAGCCAGGC
CTGCGGAACAGGGGAGAGGTTCGAAGTTGGGTAAATTCGGTCTGTGGCCTGTCTTACCCGATCGACAATAGCCAGGGCTG
CTACGTCGCCCAGTGCGACCTCCAGGGCGGCGAGCAGGTTGGCCATCAAGAACTGCTCACGCTCGTCCGCGTGGTTTTTT
GCGCTTGGGCGTCTATTCCAGTCGATCTGTGCTGAATACCTCGCTGTTACCGGGTCTGGCTCGCAGTAGGACGGCCCGAT
GTGAACCTGGCAGTCCAGGCATTCAATGGCACAGTTCCAGGCATCGGTGTCGAGCAGCTCGGGTTCCAGCATTCGAATGT
TCGTGCTGCCGCAGAGCGGGCAGGGCTTGAGCTTGGGGTCACTCATGGCCTGCCTCCCTGCCCGTCGTCGGTGCTGGGCT
TAGCTTGGGAGGAGCCGGCGACGAGCAGGTCGGCCAGGTCCGGATAGTTCGGATGGGTCCAGAAGCCGAACTGATCGCGC
TCGACCAGGTGGGCCTGAATCATAGGTTTCATGCCAGGCCTCCCACGCTCATGCCGAGCACCGCATTTGCCTCGGCGGCC
AGTTGCTGGCCTTCAGTGCTTGATGAGCTGCCGACGAAGAGCATCAGCCGTTCGATGACCCTCACAGTGTCCTGCGAGGG
TTGGGCGAACTCGGCCAGGTGATCGATGGCAGTAGCTATGTTCCAGATCGCCCTGTTCATGTCCCGCTGGATCACTTGGA
TCGAACTGGCATTGGCCGCGATGGCCAGTTGCGATTTCGTCTGCAACAGTAGGTCGTGGGCGATTTGGAGCAGTGCGTTC
TGCAGGGACTGTCCACAGGCACGCTGCTGAGCGCGAACCGGAGCGACCAGGCCCTGCAACATCGCAGCACCTTCAGCCTC
GATTTGCAGATTCTCGGAAGCACCGGCGGCAGCGCCTGCTTTGGCCATGTAGCCTTGCACGTAGGCGTCGAACTGGGTCT
CGGCGATCAAGCTCAGTAGGAGGTTGTTGCGGTTCTGGAGGTTCATATGAGGCCCTCCGGAGCCGATTGCAGAGCGTCTG
TCCAGAGCCGAAGGGTTTCCTGCTCCTCGCTGGTCAGAGCGATCTGTGCGGTGCAGTCAGCGACCTTCTGCTTGATGACG
CGCAGGCGCTCCTCGACTTCACTGCCATCGTCGTACGCGATTTCGTTAGTCAGATGGCTAATGAAGTCACATCGCAGGCT
CCCCACCTGACTAGCGAAATACTCGGCTTCACTGCGCCATCCGATGAGCTGCTTTGGACGCTCCAGCCAGCGTTTGAGTT
TTGAGTGGAGCTGGTGCGTGGAGTACAGGTGATCGAACACCCTGGCCCGGTTGTAGATCAACGCGATCTGCAGGATTTCC
AGCCGATCTTCGAGCGTATTTCCCGGAAGCGGGTAGTTCGAGGCGACCAGTTGCTTGGCCAGGCGCTCACGTAGCTCGCG
CCGCTTCTTGGCTTTTTTCTCGTAGCGATCACGCTTTTCCTTCGCATGCTCGATCTGCTTGTTCACCGAGTCGAGCAAGC
GCGCTGCGTCTTGTAGGAGCTGCTTTTGTTCGAAGCTCAAGTAGCAACCCAGCTTCCCGAACAGGTCGATACCGCCACCA
GTCAGCTTGCGTATTTCTCGCGTAGCGCCGGTCAATTTGGCCTGCTGGCTCCGAAGGTCGGCGGCGCTGCCGAACTTGTT
CAGTTCCAGCACAAGTTCCGCGTCGACGTCGCGGGCTTGGGTGATAGGAGTCTTCGCCATGTCAGCGCCCTCCCTTGACC
AGTGCTGCAACCGCCTCCCGGCGTCAGCGCCAGGAAGCGGTCAGTCCACTCGGGAAAGACCTCGATGGCCAGGTTTCGAA
TGATTTCCAGCGCGGGCGCTGATGTGCGATTGGAGGGTTGCCGCGTTTCGAGGCGGTGCGCTGGTAGCCCGCGCGATGCT
GCGTTGCGAAACACGACGGCATCCGGGACAGTCGTTTCGAGCACAGAAATGTCCTGATGCTCATCGAAGATGGCACGCAC
ATTCTCGTGAATTGCCCGGGAGTCATTGGTCTGGTCCAGGCAGTTGATGACGATCTGAACATTGGGGATCCGCATGCCGA
GACGCTCATAGGGGCGTAGGCCGTCGAGCATTTGCATGGTGCCGCGGTTGAACTCACGGGCGGTAAGCATGTTGGGTTGG
AGGGGGGAAACAACCAGGTCCGATGCAAGCACAACCATTTCGAGCAAAGCTGAGCGCGCACCCTGGGTGTCGATCAGCAC
CAGGTCGTAGCCTTCTTTCAGAGCGGGCATCAGGTTCGCCAGGCGTAGCCGGCCATCGGGCGCCTGGAGCAGTAGGTTGT
TGAGCTGATTGTTCTGGTCGTTGGAAATCACGACGTCCAGATTGGGGATAATCGTCCTGGAGATGATCCTCGCCGGGTCC
GTTATGTTGGCGGCGAGCAGGTCGTAAATGCCGCCCTGGGCAACTTCCGGCAGCTCGTAGTACGAGGATAGGGAGGGCTG
GACGGGGTCCAGATCGATGAGGAGGGTTCGTATGCCTGCATCGGCGCAAAATGCACCGAGGTTGGCGGCGGTGGTGGATT
TTCCTACACCACCCTTGGTGGATACAACCGAAGTCGCTTTCGCATTCATGTCTACAGCCTCTAATGTGGGGTCATTAGAG
GCTGCTAAACAGACTCGTTGTAACGGGAGTGGAAACCCTTCCTAGAGAAACCTACGACCAGTTGTCTTTTAACCAATTGG

Figure 2

FIGURE 2P

TCGTAGGTTCGAATCCTACACGACCCACCA

FIGURE 3

RL024

DNA sequence (SEQ ID NO: 3)

GTGGCGCTGACCGGTAATCCCCTCCTGAAATTGCTGGTCGTCCCCGTCGTGATCGGCGCCATCCTGATCGGCGTGAGCAT
GATGGGCAAGAAAGAAAGTGCGCAGTCACAAGGCGCCGCAACCCCGACGGTAACGTCGGAAGAAGCGGCAACCCTGGGCA
TCGACGGCGACACGCCCGCCGACACACTACGCACCATCGTGGCGGAAAGCCGGCAGCTCAAGGACCAGATCAGCAAGGTG
ATCCAGGAGAATGACTCGCTCAAAGCCGCCAATGAGAACCTGCAGGGCCGCCTGCGCAACATCGATCAGAACATCGAGCA
GAAGCTCAACAACACCGCCCAGGAACTGCAGCAACAGCAGGAAAACCGTAGCCAGACGATCCTGGACCAGGTACAGAAAC
GGCTCGAGAACCTAACCCACATTCCCGAGGCCGGTGACACCGACCTGCCCGTAGGATTCGGCGTGCGACCAAAGGATGGC
CAGCACTTTCAGGGAGCGGGCTCGTCTTCATCGGATATCGTCTGGATCGAGCCCCAGGACGCCCGCGCGGTTGATGCCAA
TGGCCAGCCGCTGGCCGCCGGCTCCACCACCCAACCGAGCGGATTCAGCTTCCCGACCTCCTTCGGCAATGCGGTCGATC
GCGGACAGAACGCGCTGGAGCGGATCGATGACGGGCTGCACCCCGTCGGCCAACAGCGATCTGACCTGGAAAACCGCAAG
CTCGTCCGTAAGACCTACACGCTGCCGCAGAACTCGACGCTCATGGGCTCGGTGGCCATGTTTGCGCTGATCGGTCGTGT
GCCGGTCGACGGGACGGTCAATGATCCTTACCCGTTCAAAATCCTCATCGGCCCGGACAACCTCACCGCCAACGGCATCG
AGCTGCCGGACGTCGCCGGCGCGGTAGCCAGCGGGACCGCCTCGGGCGACTGGACACTCTCCTGCGTGCGTGGGCAGATC
CGCAGCCTCACGTTCGTGTTCAACGACGGGACCGTGCGCACCTTCCCGGCGCCGGCCGAGGAGGTGAATGACAACCAGAG
CAACAACAACCAGACCGCCAGCGCCGACCAGAAAACCATCCAGGGCGGCCTCGGCTGGATCAGCGACCCCTACGGCATCC
CATGCATCGCCGGTGATCGCCGATCCAATGCCAAGGAGTACCTGGGCAATCAGAGCCTACTCACGGCTGCCGGGGCCGGC
ATTGCCAAGCTCCTGGACGCCGACGAGAACAACACCAGTACCGTCTTCAGCGGCAACGGCACCAGCTTCGGGACGACCGG
AACCAACAGCAACTCGGCCCTCAACAGCATCCTCTCCGGCGGCGTCAGCGACATCCGGCAGTGGATGAACAAGTTGTACG
GGGAGGCCTTCGCCGCCGTCTACGTGCAGCCAGGTGCGCGGGTCGCAGTGCATCTCGATCAGCAACTGGCGATCGACTAT
GAACTCAAGGGCCGCAAGGTCGATTACAGCTCTGGAGCCGCTCATGCAACAGCAGACTTGGACTAA

Protein sequence (SEQ ID NO: 127)

VALTGNPLLKLLVVPVVIGAILIGVSMMGKKESAQSQGAATPTVTSEEAATLGIDGDTPADTLRTIVAESRQLKDQISKV
IQENDSLKAANENLQGRLRNIDQNIEQKLNNTAQELQQQQENRSQTILDQVQKRLENLTHIPEAGDTDLPVGFGVRPKDG
QHFQGAGSSSSDIVWIEPQDARAVDANGQPLAAGSTTQPSGFSFPTSFGNAVDRGQNALERIDDGLHPVGQQRSDLENRK
LVRKTYTLPQNSTLMGSVAMFALIGRVPVDGTVNDPYPFKILIGPDNLTANGIELPDVAGAVASGTASGDWTLSCVRGQI
RSLTFVFNDGTVRTFPAPAEEVNDNQSNNNQTASADQKTIQGGLGWISDPYGIPCIAGDRRSNAKEYLGNQSLLTAAGAG
IAKLLDADENNTSTVFSGNGTSFGTTGTNSNSALNSILSGGVSDIRQWMNKLYGEAFAAVYVQPGARVAVHLDQQLAIDY
ELKGRKVDYSSGAAHATADLD.

FIGURE 4

RL025

DNA sequence (SEQ ID NO: 4)

ATGATCCGGAAGTCGACAGGCTCGCTCTTGCTAATGCTTGCCCTACCCACACTGGCCCACGCGGTGGAGATTCTGCGCTG
GGAGCGCATTCCGTTGGCCATTCCATTGACGGTCGGCCAGGAACGCATTGTTTTCGTCGACAGAAACGTGCGAGTTGGGG
TTCCTCGGGATCTGCAGGGCAAGCTGCGCGTCCAGAGTACCGGCGGCGCACTCTACCTGCTCGCCAACGAGCCGATTCCT
CCAGCGCGCCTGCGCCTACAGGACGCGACCAATGGCGAGCAGATGCTCATCGATATCGCCGCCACCGAAGCAACGGCCGA
CCAACAGCCGCGCGAGCCGGTCAGGATCGTCGCCGGCGAGCCAGTGGATCCGCATTATGGCCAGTCCCGGGAAGCCCAGC
CATCGGCAGCAGCGAAACAGACCGAGCACGCAGAAGCACCGAAGGCCGTGCCGCGCGAAACGCCCGTCCCCGTGGTTCTG
ACGCGCTATGCGGCGCAGATGCTCTATGCCCCGCTTCGCACGGTGGAACCGGTGGATGGCGTCGGTCAGGTGCGCGTCAA
GCGACAGCTCGACCTGACCACCCTGCTCCCCAGCCTACCCATCACGGCTACCGCCTTGGGCGCCTGGCGGCTGGACGACT
ACTACATCACGGCGGTGAAGCTGCAGAACGCCAGCGCCCAGCACCTGGCCCTGGATCCCAGGGACCTGATGGGCAATTTC
GTCGCCGCGACCTTCCAGCACCCGTACTTGGGGCCCCGGGGCGACGCCTCCGACACCACTACCGTGTATCTGGTGACGCG
CGGCCGCGGCCTTGCCGACGCGCTCCTGCCCTCCTCCATCAGCCAGATCGATCCCAAAGGAGGCCGTCGTGGCGCTGACC
GGTAA

Protein sequence (SEQ ID NO: 128)

MIRKSTGSLLLMLALPTLAHAVEILRWERIPLAIPLTVGQERIVFVDRNVRVGVPRDLQGKLRVQSTGGALYLLANEPIP
PARLRLQDATNGEQMLIDIAATEATADQQPREPVRIVAGEPVDPHYGQSREAQPSAAAKQTEHAEAPKAVPRETPVPVVL
TRYAAQMLYAPLRTVEPVDGVGQVRVKRQLDLTTLLPSLPITATALGAWRLDDYYITAVKLQNASAQHLALDPRDLMGNF
VAATFQHPYLGPRGDASDTTTVYLVTRGRGLADALLPSSISQIDPKGGRRGADR.

FIGURE 5

RL026 : DNA sequence (SEQ ID NO:5)

ATGAGTTTCAGAAAACACACTGCGCAACAGCAGGCACACATCAACACCTTCCGGTTCATCACCGGCTTCCTGTGCATGGT
CATCGTTGTGCTGGCCTACTGCGTCTGGGAAGCCCGTAAGGACCTCTGGATCCACATTCCGCCCGACTTGCGCTCAGGAA
GCACCCGGTTGTGGTGGGACATTCCGCCAGAGAGCGTCTATGCGTTCGGCCTCTACATCTTCCAGCAGGTGCAGCGTTGG
CCCAAGGACGGCGAGGTGGACTACAAGGGAAACCTGTTCCGCTACGCTGCCTACCTCACTCCCTCCTGCAAAGTCTTCCT
GGAGAAAGACTTTGAGTTTCGTCGTAACGCCGGCGAGCTCAGGGGTCGCGAGCGCACCACCTCGGAAATCCCCGGTCGAG
GCATTGGCGAGAGCAATGGCCGCGTGATCCAGCACTCGATCAATGACTGGACCGTCAACTTGGACATGGACAGCACGGAG
TATTACGCCGGCGAGAAGATCAAGCGGGCGCTGGCCCGCTACCCGTTGCACGTTATCCGCGCCGACGTCGACCCGGAAAC
CAATCCCTTCGGCCTGCAGTGGGACTGCTACTCCGACACGCCTCAACGTATCGAGCTTGAGGAGCCGGCCGCCCCCACCA
AGCGGGAGGGAGGTCTATGA

Protein sequence (SEQ ID NO: 129)

MSFRKHTAQQQAHINTFRFITGFLCMVIVVLAYCVWEARKDLWIHIPPDLRSGSTRLWWDIPPESVYAFGLYIFQQVQRW
PKDGEVDYKGNLFRYAAYLTPSCKVFLEKDFEFRRNAGELRGRERTTSEIPGRGIGESNGRVIQHSINDWTVNLDMDSTE
YYAGEKIKRALARYPLHVIRADVDPETNPFGLQWDCYSDTPQRIELEEPAAPTKREGGL.

FIGURE 6

RL027
DNA sequence: (SEQ ID NO: 6)

ATGCCCGAAGAACATCTGTTTCAGGATGGAACCCTCAGCTTCCTGCCGACCC
GTTTGAACCGGCAACCGGTAGTCATCGG
CGGCCTGACCGCAGACGAAATGTGGATCACGGTCTTCACCAGCGGAGCAGCC
GGGTTCGTTCTTGGCATCCCGGCTGCCT
TGGTCGCAGGTAACGCTGCCTGCATTCCACTGGGCGCGCTGCTGGTCGGCGC
CCTCGGCCTAGGTATCGGCAGCCGCGTC
CTGCGGCGGATGAAGCGGGGGCGGCCCGATACCTGGTTCTACCGCCAGGTGG
AGATGGCCCTCTCGCTGCGCTTTCCCGT
CTTCGGCAACCGTCGCCTGGTTACGCGCTCCGGCGCCTGGACCAGTCGACGC
ACGGAGTCCCCATGA

Protein sequence: (SEQ ID NO: 130)

MPEEHLFQDGTLSFLPTRLNRQPVVIGGLTADEMWITVFTSGAAGFVLGIPAALV
AGNAACIPLGALLVGALGLGIGSRV
LRRMKRGRPDTWFYRQVEMALSLRFPVFGNRRLVTRSGAWTSRRTESP.

FIGURE 7

RL028

DNA sequence: (SEQ ID NO: 7)

ATGCTGAAACTCACCCTCCAGAAACTGTCCGCCCTCTGCCAGAGCCTGGCCGCCATCACTTTGGCGCTCCCCGGTATCGC
CTTGGCTGCACTCCCCAAACCCGAGGCACCTAGCCGTGGGGAGGGATCGGGCATCATGCAAACCATCCAGAACTTCGGCT
ATGACGGAGCGATGCTCCTCGCGCTGCTCATCTGCGCGGCTGTCTTTCTGGGGGTCGCTTGGCATACCTACGGCACCTAT
CACGCCATCCATGACGGGAAGAAGAAGTGGTCGGATCTCGGAGCGGGCGTAGCCGTAGGTGTCGGCCTGCTGATCTTGAT
CATTTATCTCGTCACCAAAGCCACCGCCATCATGTAA

Protein sequence: (SEQ ID NO: 131)

MLKLTLQKLSALCQSLAAITLALPGIALAALPKPEAPSRGEGSGIMQTIQNFGYDGAMLLALLICAAVFLGVAWHTYGTY
HAIHDGKKKWSDLGAGVAVGVGLLILIIYLVTKATAIM

FIGURE 8

RL029

DNA sequence: (SEQ ID NO:8)

ATGAGCATGAGCGGAGCCCAGACATCAGCGTTCCAGGCCGCCGCTGGCTTTCCCCCATCGGCCGGCGAGGGACTGTTCAT
TGGAGCAGCGATGACCTTCCTTCTGCTGTGGTCCGCCTGGGCGATGTACAGCACCTGGCGCGGCTGGGCCACCAACAACC
TTCGACAGCGCCACCGGTGGCGCTTCCGCGATCCCGGATCTTGGTCCTCCTCGGCATCACCTCTTTCTTCCTCCTCAGCT
GACCCATACGGAGACACTCATGCTGAAACTCACCCTCCAGAAACTGTCCGCCCTCTGCCAGAGCCTGGCCGCCATCACTT
TGGCGCTCCCCGGTATCGCCTTGGCTGCACTCCCCAAACCCGAGGCACCTAG

Protein sequence: (SEQ ID NO: 132)

MSMSGAQTSAFQAAAGFPPSAGEGLFIGAAMTFLLLWSAWAMYSTWRGWATNNLRQRHRWRFRDPGSWSSSASPLSSSSA
DPYGDTHAETHPPETVRPLPEPGRHHFGAPRYRLGCTPQTRGT.

FIGURE 9

RL030

DNA sequence: (SEQ ID NO: 9)

CTGATCTGCACGAGATTCGCCGTGAACACTCCACATCCATCCCTTCGCCGAAGCTGCCTGGCCGTCTTGGCCTGCAGTGC
GCTGGTCGCACAGGGAGCTTTCGCAGCGAGCGCCTCCGAGCAGGCGAACCTGGAGGTGATGATCCGGCAGCTCAACGCCC
TCGAGGACACCGCCCGCCGCAGTGCCCAGGGCGCCGATGAGCCCGGACAGCGCTTCTACTTCGACTACCCGCGCCTGGCC
GCTGACCTGCAGCGCATCCGCCAAGGCCTGCAGGACTACATGACGCCCAGCCGCGCCCAACCGCGTGACCCTTCCGACTT
ATCAGGGAATTACACCCTGCGCGGAGGGCCGATGCCATGA

Protein sequence (SEQ ID NO: 133)

LICTRFAVNTPHPSLRRSCLAVLACSALVAQGAFAASASEQANLEVMIRQLNALEDTARRSAQGADEPGQRFYFDYPRLA
ADLQRIRQGLQDYMTPSRAQPRDPSDLSGNYTLRGGPMP

FIGURE 10

RL031

DNA sequence: (SEQ ID NO: 10)

ATGAGCATAAAACAGCCCTTCGAATACCATGTCGAGAACATCGTCATTCCCTACAAAACCCTCACCAAGGGCGTCGCGAT
GTTCAAACACAAAGAAGACACCTTGGAACCCGACGACCACGCCTTGCTCAACCCTCTGCGCTGGGCCGAGGTCGTGCGTC
TGGGCCAGGAAGGCTGGGAGCTGGTGAGCGTTCAGCCACTCATGCGGGGCGTAACCGAGATCGGTAATCAAAACGCCCAA
GGCTGGGCTTGGGGCGTCGCTCTGCCCGTCAGCTACCTGCTGTTTTTCAAGCGCGCAACCTCATAA

Protein sequence: (SEQ ID NO: 134)

MSIKQPFEYHVENIVIPYKTLTKGVAMFKHKEDTLEPDDHALLNPLRWAEVVRLGQEGWELVSVQPLMRGVTEIGNQNAQ
GWAWGVALPVSYLLFFKRATS.

FIGURE 11

RL032
DNA sequence: (SEQ ID NO: 11)

ATGCTTAGAAACATCTCTATTGGAGTTTTGCTAGCCATGGCTGCTATGTTGGGCAGTTATGGGGTGGCTGCCGCTACATT
ACGATGCGGGTCGGCAATTGTTAGTGAGGGCGACTTGATTGATGATGTGCTTAGAAAGTGCGGCAACCCTGATAGCCGTA
AAATTGAAGGGCCCGCAGTGGATGGTAGTGGCTATATAGTGCGGGGGGCTGCTACTGTCGAAAACTGGGTATATGGACCA
AGGAATGGATGGTACCAGAAGCTTAGGTTTGTCGATGGAAGACTAGTTCAGATAAAAGGCAGTATGGACTAG

Protein sequence: (SEQ ID NO: 135)
MLRNISIGVLLAMAAMLGSYGVAAATLRCGSAIVSEGDLIDDVLRKCGNPDSRKIEGPAVDGSGYIVRGAATVENWVYGP
RNGWYQKLRFVDGRLVQIKGSMD

FIGURE 12

RLO33
DNA sequence: (SEQ ID NO: 12)

ATGAAACTTATCCTTGATTTCGACGGACGCCTTCTAAATCCAAGCAACATGCTAGAGGCCCTATCAAAAGCAGGAAAAAA
TACAAGCATCAGCATAAGCAACGCGCAAGCATTAAATATAGAAACTCTTCTCAAGGCAACAACCACTGCAGAAAACACAA
AAAACCTCTCAACAACTTTCAACGGCGCAGAGCTGACTGCTAACAACCTTCAGCAAGTCATAAACTCAGCAGGATCACTA
ACCAGAGTATCCACAATAGCCGCACAAGCCATTAATATAAACACACTTCTTTCCGCAATATCTACAGCAGGCAACTCAAA
GAATTTTAGCGCAGAATTCAATGGAGCCCAACTCAGCAGCGACAACCTACTTAGAGCAGTAAATGCGGCAGGAACAAACA
CCAGCATAAGCGTCAATACCGCACAAGCGGCAAATATAACCGCCCTTCTTCAAACTATTCATGCAGCAGGTGACACAAAA
ACATTCAGCGCAGAGTTCAATGGCGCTCAACTTACTTCAAACAACATTCAACAAGCTTTAGACGCCGCAGGAACCCGAAC
ATCCATTAGCGTCAATACCGCACAGGCGGTTAATATAAGCACCCTACTAGCCCTCATCAACTCTGCCAAAGACACGAAAA
AGTTTAGCGCCGACTTCAATGGTGCACAACTAACAGCAGACAACCTTCAGCAAGCGATCAGCGCTGCGGCCTCGGGTACC
AATATCAGCGTCAACACCGCTCAGGCGGCGAATATATCCACCCTTTTACAGGCCATCAACATCGCGGGCAACACTAAAAA
ATTCAGCGCCAACTTTAATGGTGCCCAACTCACTTCAAACAACATCCAGCAGGCGCTCCGAGCGACAGGATCAAACACAT
CAATCAGCATGAACTCCGCACAATCCGCCAACCAAAGCACTCTACTTGAACTTCTAGACATAGCAAGTTCCAGCAAGCAA
TTCCAAGCCAATTACAACGGTGGCATGTCTAATCCGAACAACCTACAACAGATAGTTTTCCCGTGCAGGCGCCAGTACAA
CCGTGTTTATTTCCGACGCACAAGGCCTACCAATCGCAAATATCCTTACCCTTATATCATCTGCCGGATGAGACTTATAG
CCGTGGATGAAAACACACCATCCACGGCTATACCCTAG

Protein sequence: (SEQ ID NO: 136)

MKLILDFDGRLLNPSNMLEALSKAGKNTSISISNAQALNIETLLKATTTAENTKNLSTTFNGAELTANNLQQVINSAGSL
TRVSTIAAQAININTLLSAISTAGNSKNFSAEFNGAQLSSDNLLRAVNAAGTNTSISVNTAQAANITALLQTIHAAGDTK
TFSAEFNGAQLTSNNIQQALDAAGTRTSISVNTAQAVNISTLLALINSAKDTKKFSADFNGAQLTADNLQQAISAAASGT
NISVNTAQAANISTLLQAINIAGNTKKFSANFNGAQLTSNNIQQALRATGSNTSISMNSAQSANQSTLLELLDIASSSKQ

FQANYNGGMSNPNNLQQIVFPCRRQYNRVYFRRTRPTNRKYPYPYIICRMRLIAVDENTPSTAIP.

FIGURE 13

RLO34

DNA sequence: (SEQ ID NO: 13)

GTGCAGTGGACTCACGAACAGTCACCGATCATCCAGTCGAAGGCACCGAAGATCCTGGTGCGAGCCTTCGCAGGCACTGG
CAAAACTACCACCCTGGTGGGCTTTGCCAGGTCGAACCCTACCCTGAGAATCCTCTATCTCTGCTACAACAGCTCGGTGG
AGAAAGCCGCGAAGGGCAAGTTTCCCCGCAACGTAGTGTGCAAGACCGCCCACAGTCTGGCTCATGCGGTGTACGGCATT
CAGTACGCCCACAAGAAGACGAAGAACCTGCGACTGACCGATATCGCCCGCGGACTCGATACCCAAGACTGGGAGTTGGT
ACGTGACGTGCTGGCCACGCTGAACAACTACATGGCCAGCGCCGACGCGGAACTCGGCCGACCGCACTTCCCGCGCTTCC
GCGACAAGGCGTTCCTCACCAGTGCTCAGGAACGCTTCCTCAAGCAGGGCCTGGACATGGCGCGAGTAGTCTGGAGGCGC
ATGGTCGATCTCCAGGACACCGGCATGCTGATGCCCCTTGACGGCTACCTGAAGCTGTATCAACTGAGCAAGCCCGATTT
GAGCCAGCGCTTCGACTGCATGCTCCTGGACGAGGGGCAGGACATCAACCCAGTGATCGCGGACATTGCCCATTGGCAGC
GCATCAGAATGGCTATCGTCGGCGATCCCCATCAGCAGCTCTACCGGTTCAGGGGCGCAGAAGATGCCCTGAACAGCGAC
TGGATGGCCGGCGCCGAGGAGCACTACCTGACCCAGAGCTGGCGATTCGGCCCCGCGATCGCACACGTGGCCAACATCAT
CCTCTCCTACAAGGGCGAAACACGGAAACTTCAAGGACTGGGTCCGCAGACGCTGGTGAAAAAGTCCCTCCCGCCGGACC
TTCCTCACCGCACTTTCATTCACCGCACCGTTATCGGCGTCATCGAGAATGCCCTGCAGCTGGTCCGCAATCATCCGGAG
CCCAAATTCCACTGGGTAGGCGGTATCGACAGTTACTCGCTGCGCGACCTGGAGGATCTGTACGCATTCAGCCGAGGCCT
GCGCCAAAACGTCCAGAACAAGAAACTGCTCCGTGACTACCGCGACTACACCCAGTACGTGGAGATCGCCGAGATCAGCC
AGGACGGTGAGATGCTTCGCTCGATCAAGATCATATCGACCTACCCTGATCTGCCTGCGCGGATCCTTGAGCTTCGCTCA
CTGACCCTTGACGATGAGCTGGACGCAACAATCACCCTGACCACCGCACACAAGGCCAAGGGGCTGGAATGGGATTTCGT
TTGCCTGTACGACGACTTCAACGCGGACCCGCTGGCCCCCGACACCGACCCAGGCAAGCGCGACGATGAGTTGAACCTGA
TCTACGTCGCAGTGACCCGCGCGATGAAGATCCTTGCCATCAACAGCCTGGTGCTGTCGATCATGCAGCGGTACGTGGAC
GACAGAAAACTGAAGGAGCAGATAGCTAGCTGTAAAAAATGA

Protein sequence: (SEQ ID NO: 137)

VQWTHEQSPIIQSKAPKILVRAFAGTGKTTTLVGFARSNPTLRILYLCYNSSVEKAAKGKFPRNVVCKTAHSLAHAVYGI
QYAHKKTKNLRLTDIARGLDTQDWELVRDVLATLNNYMASADAELGRPHFPRFRDKAFLTSAQERFLKQGLDMARVVWRR
MVDLQDTGMLMPLDGYLKLYQLSKPDLSQRFDCMLLDEGQDINPVIADIAHWQRIRMAIVGDPHQQLYRFRGAEDALNSD
WMAGAEEHYLTQSWRFGPAIAHVANIILSYKGETRKLQGLGPQTLVKKSLPPDLPHRTFIHRTVIGVIENALQLVRNHPE
PKFHWVGGIDSYSLRDLEDLYAFSRGLRQNVQNKKLLRDYRDYTQYVEIAEISQDGEMLRSIKIISTYPDLPARILELRS
LTLDDELDATITLTTAHKAKGLEWDFVCLYDDFNADPLAPDTDPGKRDDELNLIYVAVTRAMKILAINSLVLSIMQRYVD
DRKLKEQIASCKK.

FIGURE 14

RLO35
DNA sequence: (SEQ ID NO: 14)

ATGTTCGGGTCGCTGATCGGCGCAATCATCGTGGAGTGGGTATGCCTGTATTTCTTCTGGCCTGACGCGGGCTGGAAGCA
TGCCCAGGCCATGTTTGAGTACGAACTCAGTTGGCTGTCGCAGGGGCTGCTACACAGCGTCGTCGTGCAGGAGCCAGGTC
GAACCGCCACCTGGCTGGCCCAGTTGGCCTATGACTGGTTGTTCGTGAAGACCGGGATGGTCGACTGGATGACCAACATG
ACTACCATCGCGCAGGCCCGGCCACGGAGCCCGCTGGACGTTCGCTATCTCACCGCCCACGGTGTCTCCACGCTGCAGAA
CTACGGCCTGGCCGCGCTGTACACGGTGCTGACATTCGTCGTGCGCCTGGTGATCCTGGTCATGACGATCCCGTTATTCG
TGATGGCCGCGTTCACCGGCCTGGTGGACGGCCTGGTGCGCCGGGACCTGCGCAAGTTCGGCGCCGGCCGGGAGTCCAGC
TACCTCTACCACAAGGCGCGCGGCAGCATCATTCCGCTAGCGGTCGTCCCTTGGACGCTCTACCTGGCAATCCCCATCAG
CATCAATCCCCTGCTCATCCTGTTGCCCTGCGCCGCGCTGCTCGGCGTAGCGGTATGCATCACAGCATCCACCTTCAAAA
AGTACCTGTAG

Protein sequence: (SEQ ID NO: 138)

MFGSLIGAIIVEWVCLYFFWPDAGWKHAQAMFEYELSWLSQGLLHSVVVQEPGRTATWLAQLAYDWLFVKTGMVDWMTNM
TTIAQARPRSPLDVRYLTAHGVSTLQNYGLAALYTVLTFVVRLVILVMTIPLFVMAAFTGLVDGLVRRDLRKFGAGRESS
YLYHKARGSIIPLAVVPWTLYLAIPISINPLLILLPCAALLGVAVCITASTFKKYL.

FIGURE 15

RL036
DNA sequence: (SEQ ID NO: 15)

ATGAAGTTGAAGAATTTCTTACAGCCTTTTGATAGCGGTTTCTCCACTCCGAGTGCTGCGCTCAAGCTGCTCCGCATGCT
CGGTGGCGCCTTGATGTTGTGCGTGCTATGCAGCCTGATATTCAGTGTGAGCATGGTTTTAAACCATCAGGTGTCCCTCA
GTCGGCAAGCTATGAATGTGGCTATGTACGAAGCGCAGCTTTATTTCGAGCAGCGCGAGGCGTTGCTCAATCACTTGAGC
GGCAATGTCGTGCCCTTGGCCGCGGGTAGAGCGCTCGTCAACGAAGCGCCGAACAATGTGAGCATCCTGCCGTTGAGTGA
CGGAGGGCGAGGTCTGCTATTGACCGCTCGCACGCTCGGTGATCTCCGGGAAAAGCGGCTGGCACTGATGTATCTGGTCG
ATACCGACAAAGGCCCTCTGGTTTACCGGCTTACCGCCGATGGTAGGCCCTCGGCAGCGATATCCAGCACGATAACCAAA
GAGGTGTACCGAGCCTTGCTGGCGACTCCGTCGGCGCCTGTTCACTGGGTGACTGACGGTGGTACCCCTCAACGGCTGTA
CCTTTTTGAATCCTTAGGCGATGAGCCGGGCGAGGGGTGGCTAGGCCTGGAGATTCTCGGCGAAGACCTCGATTCGATGT
TGCGCCGGAATGATGCCGGAAACTACATGCTGCTGGATCAGCATGGGCAGGTCGTACTCGCTACGGACGCAGAGGCGCTG
GGGAGCGGTGCGTCGCGGACGCTTTTGCGTGGAGACGGCTTCGGTTTCATCGGTGCTGGCCCACTGCCGCAGCATATGGT
GCTTTTCCAGCACGTGGGGTCTTCGAGCTGGGATCTGATCTATCACATCGGTATCGGTCGCCTGTTGCTGGCTCTGTGGC
TCCCTCTGTTACTTGCCTCTGCGTTGGCACTCGCAGTCGGCATCCTACTGCATTGGCTGGTGCGGAGCATCGAGCGACGC
TTGATAGAGCCCGCAAAGCGACGCCTTGAAGCATTGAAGGAGAGCGAAGCCTTTTCCCGTGCAGTTATCCAGGCCGCGCC
CGTCGCGCTGTGCGTGCTGCGTCGTGCCGACGCCGCAGTGGTCCTGGAAAATCCCCAGGCGCGCCAATGGCTGGGTGATA
GCGAGGCGATTGCCCACGACGCGCCGAGATGGATTTCCCAGGCGTTCGCAGGAGGTGTGAAGTGTTCTGGAGAAGAACTG
GAAACCGAGGCAGGGCTACATCTTCATCTCAATTACACGCCCACCCGCTATAACGGTGAAGACGTATTGTTCTGCGCCTT
CAGTGAAATCAGTGCACGCAAGCGGATGGAGGCGGAACTGGCTCGCGCAAAATCCCTGGCGGATGCTGCCAATGAAGCCA
AGACGCTGTTTCTCGCCACCATGAGCCATGAAATCCGCACACCTCTGTACGGCATGCTTGGCACGCTTGAGCTGCTTGGG
CGTACCGAGCTGAGTCGGCAGCAGGCCGGTTACCTAAAGGCAATCCAGCATTCCTCGTCGACCCTGCTGCAACTGATCAG
CGATGTGCTTGACGTATCCAAGATAGAGGCCGGCCAACTGGACCTAGAGTGCGTGGAATTCTCCCCGCTGGAATTGACCG
AAGAGGTCGTGCAGTCGTTCACCGGTGCCGCGCAGGCCAAGGGGCTGCAGTTGTATACCTGCCTCTCTGCGGAGCTGCCG
CTGCGCATGCGGGGGGCCGCGGCGTCGATCCGGCAGATTCTCAACAACCTGCTGAGCAACGCGGTGAAGTTCACCGACAA
TGGCTATGTCAACGTCCACCTGAAGGCCAGCGTGGTCGATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCG
GCATGGGGATCAACGTCGAGGATCAGCCGCGTCTGTTCGAACCGTTCTACCAGATACGCCGCTCCGAGCATCCGGTCGCA
GGCACGGGCCTCGGCTTGTCGATCAGCCAGCGCCTGGCGCAGCTAATGAATGGCAGTCTGAAACTGGTCAGTGAGCTGGG
GTTGGGCAGCAGCTTTAGCCTCAGGCTTCCGCTTGAGCGGATCGCGATGCAGGCTGAGCCGCAGGACCTAGCCGGGTGCG
CCGTCCAAGTGCTGGCGCCTGTCCGCGACCTAACGGAATGCCTGTGTGGCTGGATCTCCCGCTGGGGTGGAAGGGCCATG
GTCGCGACGCCGAGGTCGCTGGACGAGGCGGACGCGACCTCGCTGCTGGTCAAAGTGTTACTGCTGGAGGGGGCGCCGAT
GTTCGAAGCATGGCCAGGATGCCGGGTGGAGCTTTCCCCTCAGGGTGATATGGAGCCGCAGGCACAGGGCCGCGACTGGC
TGCTCGGGCTCAACAACCTGAACGGCCTGCATCGTGCTCTGGGCCTGGCCCATGGGCGTCTCGCTGATCCTTCGACGCCG
CCGATACGGCTGGCTCCGTTGCGCAATCTAGGTCTCCGCGTCCTAGTGGTGGAGGATAACGCGATCAACCAGTTGATCTT
GAGGGACCAGATGGAAGCGCTGGGCTGCAGCGTGGAGCTGCTCTTCGATGGTCGCGAGGCGTTGCTGCACTGCCAGACGG
CCTGCTTCGACGTGGTGCTCACCGATATCAACATGCCGAACATGAACGGATACGAGCTAACCGCGGAGCTACGGCGCCAA
GGGTTCCGGCAGCCGATCATCGGCGCGACGGTGAACGCCATGCGTGAGGAGCGCGAGCGCTGCATGTCCGCCGGGATGAA
CGATTGCCTGGTCAAACCGGTGGATCTGAATGCCCTTCAGAACTGCTTGATTAATATTCTCAAGGTGGATCGATGA

Protein sequence: (SEQ ID NO: 139)

MKLKNFLQPFDSGFSTPSAALKLLRMLGGALMLCVLCSLIFSVSMVLNHQVSLSRQAMNVAMYEAQLYFEQREALLNHLS
GNVVPLAAGRALVNEAPNNVSILPLSDGGRGLLLTARTLGDLREKRLALMYLVDTDKGPLVYRLTADGRPSAAISSTITK
EVYRALLATPSAPVHWVTDGGTPQRLYLFESLGDEPGEGWLGLEILGEDLDSMLRRNDAGNYMLLDQHGQVVLATDAEAL
GSGASRTLLRGDGFGFIGAGPLPQHMVLFQHVGSSSWDLIYHIGIGRLLLALWLPLLLASALALAVGILLHWLVRSIERR
LIEPAKRRLEALKESEAFSRAVIQAAPVALCVLRRADAAVVLENPQARQWLGDSEAIAHDAPRWISQAFAGGVKCSGEEL
ETEAGLHLHLNYTPTRYNGEDVLFCAFSEISARKRMEAELARAKSLADAANEAKTLFLATMSHEIRTPLYGMLGTLELLG
RTELSRQQAGYLKAIQHSSSTLLQLISDVLDVSKIEAGQLDLECVEFSPLELTEEVVQSFTGAAQAKGLQLYTCLSAELP
LRMRGAAASIRQILNNLLSNAVKFTDNGYVNVHLKASVVDAECVMLTWQVNDTGMGINVEDQPRLFEPFYQIRRSEHPVA
GTGLGLSISQRLAQLMNGSLKLVSELGLGSSFSLRLPLERIAMQAEPQDLAGCAVQVLAPVRDLTECLCGWISRWGGRAM
VATPRSLDEADATSLLVKVLLLEGAPMFEAWPGCRVELSPQGDMEPQAQGRDWLLGLNNLNGLHRALGLAHGRLADPSTP
PIRLAPLRNLGLRVLVVEDNAINQLILRDQMEALGCSVELLFDGREALLHCQTACFDVVLTDINMPNMNGYELTAELRRQ
GFRQPIIGATVNAMREERERCMSAGMNDCLVKPVDLNALQNCLINILKVDR.

FIGURE 16

RL037

DNA sequence: (SEQ ID NO: 16)

ATGAGCTGGAAATCCTATCGGGTGCTGGTGGTCGAAGATCAGCCGTTTCAGCGCGAATACCTGCTCAACCTGTTTCGCGA
GCGCGGCGTGCAGTACCTGGTAGGTGCCGGCGACGGCGCGGAGGCGTTGCGCTGCCTGAAGCAGGACAGGTTCGACCTGA
TCCTCAGCGATCTGATGATGCCGGGCATGGATGGTATCCAAATGATCCTGCAACTGCCGTATCTCAAGCATCGTCCGAAG
CTGGCGCTGATGAGCTCCTCGTCGCAGCGGATGATGCTCAGTGCCAGCCGGGTCGCCCAGAGTCTCGGCTTGTCGGTAAT
CGACCTGTTGCCCAAGCCGACTCTGCCCAAGGCCATCGGCCAACTTCTGGAACACCTGGAAAGATGCCTCAGGCAGAAGC
TGGAGCCGGAAACCGACGAGACTCCGCATGGGCGCACGGCGTTGCTGGATGCCCTGCATAACGAGCAACTGGTGACCTGG
TTCCAGGCTAAGAAATCCCTCCACACCGGGCGCATAGTCGGCGCCGAGGCGTTGATACGCTGGAGCCACCCGCAGCATGG
CCTGTTGCTGCCCAGCTGTTTCATGAGTGATGTCGACGCTACCGGTCTGCACGAGGCGTTGCTCTGGCGCGTGCTCGAAC
AGACCCTGAACGCCCAGGAATCGTGGCGCAGGGCGGGTTACGAGATTCCGGTTTCGGTGAATCTGCCGCCGCACCTGCTC
GATAACCAGGAACTTCCGGATCGACTCTATGAGTACGTCGGCGCTCGCGGGGCTTGTACCAGCTCACTATGTTTCGAGTT
GACCGAGAGCAGTGTCACAACTCTGTCAAGTAACTACTATGCAGGTGCCTGTCGCTTGCGCATGAAAGGGTTCGGATTGG
CCCAGGACGACTTTGGCCAGGGTTACAGCTCGTTCTATAACCTGGTCACGACGCCTTTCACGGAGCTGAAGATCGACCGC
TCCCTAGTCCAGGGATGCGTAGAGGATAACGGCCTCAATGCAGCTGTCATCAGTTGTATTGAGTTGGGTCACCGCCTGAA
TCTCGACGTGGTGGCCGAAGGCGTGGAGACCTGCGAGGAACTGAATCTTCTTCGTCGTCTTGGCTGCGACCGGGCGCAGG
GTTTCCTGATTTCTAAGGCAGTGTCTGCTCGTGAGTTCGAGCGGCAGTTAAGGGAGGACGGCCCCAGCCTCCTTGTTTAA

Protein sequence: (SEQ ID NO: 140)

MSWKSYRVLVVEDQPFQREYLLNLFRERGVQYLVGAGDGAEALRCLKQDRFDLILSDLMMPGMDGIQMILQLPYLKHRPK
LALMSSSSQRMMLSASRVAQSLGLSVIDLLPKPTLPKAIGQLLEHLERCLRQKLEPETDETPHGRTALLDALHNEQLVTW
FQAKKSLHTGRIVGAEALIRWSHPQHGLLLPSCFMSDVDATGLHEALLWRVLEQTLNAQESWRRAGYEIPVSVNLPPHLL
DNQELPDRLYEYVGARGACTSSLCFELTESSVTTLSSNYYAGACRLRMKGFGLAQDDFGQGYSSFYNLVTTPFTELKIDR
SLVQGCVEDNGLNAAVISCIELGHRLNLDVVAEGVETCEELNLLRRLGCDRAQGFLISKAVSAREFERQLREDGPSLLV.

FIGURE 17

RL038

DNA sequence: (SEQ ID NO: 17)

GTGAAGTCTGCTAGTGCCTTGGAGCACGACAACAAACTTTTGCTCAAATGGACAACCCTCTCGCAGAGCCTGAGCATCGG
CTTGATCTGTGTGGTGGTGCTGACCGTATTGCTGTTCAGCATCTGTTACTGGTCGCTGGGGAGATTGTTTCAGGAGGAGG
AGGACAAAGTCTCCTTCCACTTCACCCGTATGATGGATGTTATACGGGAGCATGAGGTATTTCTTGGGCGCATCGCTCGA
AAAAGCGACAAGACCACCCAGAAGTACGACTATGACGTGGTGCCTTTGCAGCGGCACTTGTTGGCAAAGGAAAACGGATT
AGCGGTCTATGAGGGACGGGAGTTTTCCTTTGCTATGCCATTTCTACTGGCTACCAAGCACGCGTTGAGCGCCGATTCCT
CGGGAGATCCGTTTTCGCTCGGTGTATTGCTCGCCAATTTCTACGGAAGCTTCTGGAGTGTTTCCGCCTATCCCGCGCCA
CAGTTACTGATCTTTGATCTTTCCGGCAGCACCCGCCTGGCAGTGCCGTCGATTCCCTCCACAGCGCAGCGTGACAGGTT
GAGCGGAAGCTATCCGATGATAGTCGAGCGCATTCTGGCGCGCTTGCGCACCCGGCCGGTGGGGGAGGACGCTCAGCGTG
TCCATTGGATACGCGCTGATCGCTATCGCGACTCGGCGCTGGAGATGTTGGGAGTCGCCCGGGTTGATCTGCCGGAAACA
CTCTGGTGGCACGACGAGCCGAACCATCTGATCATCGCTGCGAGCCTGCTTGATCTCAGGCGAATCAATGACTTCGAACA
GTTGGTTGAGCGCCCGGCATTCGATTCGTACAGCCTGGTATCGCCGGATGGCGAGGTATTGCTCGGCGCGGCCCCTGCGA
CCGGCCTGAGGGATGGCCTGAACCTCACCCGACAGGGGGTCGCCGTTCAACTGCTCAGCCAGCCTGAGAACGGTTGGCTC
GCGGTCTACCGAACCGACTACGGCAATTTCTTTCGCCACTCCCGGTGGCTGGTGGCAGGTCTGCTGCTGACCCCGGCGCT
GGTCCTGGCGGGTTGGCTCGGGATGCGTTGGTACACCAGCAGCGTCGTCAACCCGGTGCATCGGGCGCACCGGCAACTGG
TGGAGAGCGACACCTTCAGCCGGACGCTGATACAGACCGCGCCGGTGGCTCTGGTGGTGCTGACCCAGGATGACCAGCAA
CTGGTGACCTGCAACCACTTGGCCGCCCAGTGGCTGGGCGGGCCCACGGAGATCCTTGGGCTGACTTCCAACTGGAAGCT
TTTCGATGCGCGTGGGCAGGTACCAGGAGACATCTGTATCCAGGTCGGTGGGCGCTATTTGCAGACCGCCTTCGCGGCGA
CCCGCTATGCCGGCACCGAGGCGGTACTGTGCGTATTCAACGACATCACGGTCCACTGCGAGGCGGAGACCGCGCTGTCC
AATGCGAAGCGAGCAGCGGATGCCGCCAGCCAGGCCAAGACCCTGTTCCTGGCCCGCATGAGCCATGAAATCCGTACTCC
CCTGTACGGTGTCCTTGGCACCCTGGAGTTGCTCGACCTGACCACCCTGAACGAGCGGCAACGCGCCTACCTACGCACCA
TCCAGAGTTCGTCTGCGACGCTCATGCAACTGATTAGCGATGTGCTGGATGTCTCGAAGATCGAAGCGGGGCAGATGGCT
CTGACCCTGGCCGCCTTCAATCCGCTGGACCTAGTGCGGGAAGTGCTTGGCAACTTTGCCGCCAGCGCCATGGCCAAGGA
CCTGCAGTTCTATGCCTGCATCGACACCGAAGTGCCGGCGCAACTGATCGGTGACGTGACGCGGATTCGCCAGGTGCTCA
ATAACTTGGTGAATAACGCGCTGAAGTTCACCGATATCGGACGGGTGGTCCTGCGCGTGAAGTTGCTCTCCCGCAATGAT
GGTCGAGCCCTGTTGCAGTGGCAGGTCGCCGACACCGGTATCGGTATCGCACACGAACAGCAGGAGCGCTTGTTCGAGGC
GTTCTACCAGGTTTCGGGAGCGCACCATGCCGGCGGCACGGGGCTAGGACTGTCGATCTGCTGGCATCTGGCGGAAATGA
TGGGCGGTCACCTGCGAATGGTCAGCGAGACAGGGCTCGGCAGCAGCTTCAGCCTGGTGCTCGAGTTGCCCGAGGACGAA
CAGTCCGGGCTGGCTTGCCGGCCGGGGCTCTTGAAATCCGCTTGCGTCCATGTGCGCTCGCCCGTGCGGGAGCTAGCCGA
CAGCGTAGGGGCGTGGCTGAAAGCCTGGGGCTGCAAGGTCAGCAGCGGCGAGGCGGCGCCCTCCGAGCTGGAGACTTGTG
TGCTTCTGGAGCTGCTGCCGATGGCGGCCGGGCCTGCTTCTTCGCCCTGGCCAGGCCCCCGGGTGCGCGCGTCCATGGAT
GCGCCTTGCCAGCCGGAGCTGCGTGAGGACGGCTGGCGTGTCGGCCTGCACAACCTGGCGGGAATCGGCCAGGCCCTGGC
GCAGGCTCTGGGTGGCGATATCCCCGAGCAAACGCCGGCAAATGCCTGCGCCCGCTCGGGGAGACTCGACCTGGAAGTGC
TGGTCGCCGAGGACAACCCAGTCAACCAGGCGCTGCTTCGCGAGCAACTGGAAGAGCTGGGTTGTCGCGTGAGCCTTGCC
GGCGATGGGCGGCAGGCCCTGCAGCTGTTCGACAGTGGTCGCTTCGACCTCCTGCTCAGCGACGTCAACATGCCGAACAT
GACCGGCTACGAACTGACCCAGGCGCTGCGCGAACGAGGCGAGACGCTGCCGATCATCGGCGTGACCGCCAACGCCCTGC
GAGAAGAGGGCGAGCGCTGCCGGGCAGTGGGAATGAACAGTTGGCTGGTGAAGCCGATCACTCTGCATACCTTGCATGAA
CTGCTCAGTGAGTTCGCTCGCGCAGGTGTCGTGCTTCCCGCGCAAGCGCGAGACCTCGGCCCGCCCGCGCAGCTCGACGA
CGGTCTCTCACCGCAGGTGCCGGAACGCATGCGCGCGCTTTTCCTTGAGACCATGGGCAAGGACCTGGAGGCCGCCCGGC
AAGCGATTCGCCGCAACGACCCGAAGGGGCTGCAGCAGGACCTGCATCGCATGGCCGGCTCCCTGGCGGTGATGCGTGCG
CGAACGCTGGTGGTGATGTGTCAGGGCGCCGAGGAAGGCCTGCTGGAGTCGCGCCTTGAATGTTCCGCCGTGGAGATTGG
CGAGGTGCTCGTTCATATCGAGCAGGCGCTGGAGTTTGTGAGAAAGACGGGCTGA

Protein sequence: (SEQ ID NO: 141)

VKSASALEHDNKLLLKWTTLSQSLSIGLICVVVLTVLLFSICYWSLGRLFQEEEDKVSFHFTRMMDVIREHEVFLGRIAR
KSDKTTQKYDYDVVPLQRHLLAKENGLAVYEGREFSFAMPFLLATKHALSADSSGDPFSLGVLLANFYGSFWSVSAYPAP
QLLIFDLSGSTRLAVPSIPSTAQRDRLSGSYPMIVERILARLRTRPVGEDAQRVHWIRADRYRDSALEMLGVARVDLPET
LWWHDEPNHLIIAASLLDLRRINDFEQLVERPAFDSYSLVSPDGEVLLGAAPATGLRDGLNLTRQGVAVQLLSQPENGWL
AVYRTDYGNFFRHSRWLVAGLLLTPALLLAGWLGMRWYTSSVVNPVHRAHRQLVESDTFSRTLIQTAPVALVVLTQDDQQ
LVTCNHLAAQWLGGPTEILGLTSNWKLFDARGQVPGDICIQVGGRYLQTAFAATRYAGTEAVLCVFNDITVHCEAETALS
NAKRAADAASQAKTLFLARMSHEIRTPLYGVLGTLELLDLTTLNERQRAYLRTIQSSSATLMQLISDVLDVSKIEAGQMA
LTLAAFNPLDLVREVLGNFAASAMAKDLQFYACIDTEVPAQLIGDVTRIRQVLNNLVNNALKFTDIGRVVLRVKLLSRND
GRALLQWQVADTGIGIAHEQQERLFEAFYQVSGAHHAGGTGLGLSICWHLAEMMGGHLRMVSETGLGSSFSLVLELPEDE
QSGLACRPGLLKSACVHVRSPVRELADSVGAWLKAWGCKVSSGEAAPSELETCVLLELLPMAAGPASSPWPGPRVRASMD
APCQPELREDGWRVGLHNLAGIGQALAQALGGDIPEQTPANACARSGRLDLEVLVAEDNPVNQALLREQLEELGCRVSLA
GDGRQALQLFDSGRFDLLLSDVNMPNMTGYELTQALRERGETLPIIGVTANALREEGERCRAVGMNSWLVKPITLHTLHE
LLSEFARAGVVLPAQARDLGPPAQLDDGLSPQVPERMRALFLETMGKDLEAARQAIRRNDPKGLQQDLHRMAGSLAVMRA
RTLVVMCQGAEEGLLESRLECSAVEIGEVLVHIEQALEFVRKTG.

FIGURE 18

RL039

DNA sequence: (SEQ ID NO: 18)

ATGCGTCCGGGGTCAATAGTTGGAATTAGAACACAAGAGAAGCCTATGAGTAAGCTCAAGATAGTACTGGCCGATGACCA
TCCGATCGTGCGTATGGGCGTATGCGACATGCTCGAGCGCGACGGTCGGTTCGAGGTGGTGGGCGAGGCCTCCACGCCCA
GCGAACTGGTCGAGGTGTGCCGGCAGAGCGAGCCGCATATCGCCATTACCGACTACAGCATGCCCGGGGACGAGCGCTAC
GGCGATGGACTGAAACTGATCGACTACCTGTTGCGCAACTTTCCTCGTACTAAGGTGCTCATCTTCACCATGGTCGGCAA
CCGCCTGATCCTCGACAGCCTCTACGATCACGGGGTGTCCGGCGTGGTGCTGAAGAGCGGCGAACTCGACGAGCTGCTCT
TGGCGCTCGACGTGGTGAAGCAGAACCGCGTCTACCGGGGCGCGAACATGCTCGACCCGACCAGTGTTCTGGCGAACCGC
GACGAAGTGGAAAGCCGCTTCGCGCGCTTGTCGATGAAGGAGTTCGAAGTACTCCGTCACTTCGTTTCCGGCAGCAACGT
CTGCGATATCGCACGGCTGCTGAAACGTAGCGTCAAGACCGTAAGCACGCAGAAGGTCTCGGCGATGCGCAAGCTGGAAG
TGAACAGCGACCAGGCCTTGATGACCTTCTGCGTGCATGCCAACTTGTTCCATTGA

Protein sequence: (SEQ ID NO: 142)

MRPGSIVGIRTQEKPMSKLKIVLADDHPIVRMGVCDMLERDGRFEVVGEASTPSELVEVCRQSEPHIAITDYSMPGDERY
GDGLKLIDYLLRNFPRTKVLIFTMVGNRLILDSLYDHGVSGVVLKSGELDELLLALDVVKQNRVYRGANMLDPTSVLANR
DEVESRFARLSMKEFEVLRHFVSGSNVCDIARLLKRSVKTVSTQKVSAMRKLEVNSDQALMTFCVHANLFH.

FIGURE 19

RL040

DNA sequence: (SEQ ID NO: 19)

```
GTGTCCAGTAAGATCCTGCTGCAAGGGGCACTGCTCGGCCTAGCAATGCTGGCCGTGCTGGACGCCCGAGCCGGAGTCAC
CGCCGAGCGCACTCGGGCAATAATCGCCGAGGGGCACCGCGAGACGTCGCTGCTGCTGGTCAACCAGAATGCCTATCCGG
TCATAGTGCAGACCTGGATCGACGATGGCGCCCCGAACTCGACACCGCAGTCTGCCCGCGCGCCGATCATGCCGCTACCG
CCGGTGTTCCGCCTCGAACCCGGACAGCAACGCAGCCTGCGCCTGCTGCGGACCGGCCAGGCGCTGCCAGGGGACCGCGA
ATCGCTGTACTGGTTGAACCTCTACGAAATCCCGCCGCAAGCCACCGGGCTGCTGGCCGAAGGACAGTCACGGCTGACCG
TTACACTGCGCACCCAGATGAAAGTCATCTACCGCCCTCGCCCTCTGGCCAGAGGTGCGGAAGAAGCGCCACACCAGCTC
AGGTTCGAGCGGCGGGGCGAAACACTACAGATGGAGAACCCTACTCCCTATTTCATCAGCCTCGCCGGCGCCGAGCTTGG
CGGCCACACCCGCCTGGCGGCGGCCGAACTGTTGCCCCCCTTCTCCAGGCGCGTCCTGGCGCTCCGCCAGGCGCTGCCCG
GCGGCCAGGCCGAGGTGCGCTTCAGCTGGATCGATGACGGCGGCAATCTCCAGCAGGGACGGAGCCTGCTTCACTGA
```

Protein sequence: (SEQ ID NO: 143)

```
VSSKILLQGALLGLAMLAVLDARAGVTAERTRAIIAEGHRETSLLLVNQNAYPVIVQTWIDDGAPNSTPQSARAPIMPLP
PVFRLEPGQQRSLRLLRTGQALPGDRESLYWLNLYEIPPQATGLLAEGQSRLTVTLRTQMKVIYRPRPLARGAEEAPHQL
RFERRGETLQMENPTPYFISLAGAELGGHTRLAAAELLPPFSRRVLALRQALPGGQAEVRFSWIDDGGNLQQGRSLLH
```

FIGURE 20

RL041
DNA sequence: (SEQ ID NO: 20)

ATGAAAACATCCCTGCGCGTCCTGCCTCTGCTCCTCGCGCTGCTCGCCTCGTCTAGCTGGGCGACCTGCTACAAGGTCAC
GGCGGTAGGCAACGCCACGACTACCTCCAACACCCAGATACGTCCCGGTGAAGGCTCTGCCGGCACCTGGGCCGGAGCCT
GCGATACCTGCAACGGTTCCCTCGGTCTACCGAGCGTGATCAACGTCAGCGACGCCAGCTTCCAGCCCGACGGTAGCTTG
ATCGCCAGCTCGGTGGCGCCGCTCAGCCAATACGGCGACAGCGCCGGCTACGACCCAGAGCGCGTGTTCTTCCGCTGTGC
TCCAGAGGACGATGTGTACGAGATGTTCTCCACCAATGCCGACGATCTCTACAGCGGCTGGTACCTAGGAGGCGACAGTG
CGGGCAACTCGATTGGCCTGCAGTCCGCCTATCGCACCGCCTGGCCCAACGTGCTGCTGCGCCTAACCCACGTGGAAACC
GGGCAGTATTTCACCGATGTCTGGCGCGAGCGTCTGCTCGGCGGGCTCGATATCGACTCGCGAGGCTTTCAACTGGTCAA
GGCGAAGAACCTCAGCGCGGTACGCGCCGAACTGTTCCGCGCTCCGCTGGAGTTCATCCGCTACTACTCGCCGACTACCG
CCTCGCGGTTGTACGCCTACACCCAGCCCGCTGGCTACATCGCCATCAAGGGTCCCGGCCTGGCCTACCCCAACGTCGGC
GCCAGCCATAACGCCAACTACCTCGGCTGGCACTACAACTGGCCGGGCGCCATCGGCCTGTACAACGACGTGACGCTCAA
GCGCTATCCCACCTGTTCCGTAACCAACGTCACGCCCCACGTTGTGTTCCCGTCGATTTCCCTCAGTGAGATTAATGCCG
GCGCGAACCGTGAGATGCCCTTCGAGGTGGCCTTCAAGTGCCAAACGGGAGTGATCAACAGCACCGCCTCCAGCGGTACT
GCACTGGGTATCAGGGCTTCAGCCGGGGCGCAGGCCGCGTCCGCTGCACTGGGCCTGAGGAACGCCAATGGCGGGCTCTC
CTACCTAGTTTCCGACCGCTACGGCCAGCCTGGTATGGCCCAAGGCGTGGGTATCCGCTTGCTGCGCGACGGCAGTGCGA
TGAACCTGCTGGTAAGCGAGGATTCCGCGATGGGCAGCAATGCCGAAACACGGGGCTGGTATCCAGTGATCGGCAACGCC
TCGAACAAGACTGGCGAAGCGGGAGGCATCAGCCAGTACAGCGAGACCTTCCGTGCGCGCCTGGAAAAACTCACCGTTGG
CAGCATGCCCAGCGTTACCCCGGGACGGGTGGAGGCCAGCGCGCAGGTAGTGATTCGTGTCCAGTAA

Protein sequence: (SEQ ID NO: 144)

MKTSLRVLPLLLALLASSSWATCYKVTAVGNATTTSNTQIRPGEGSAGTWAGACDTCNGSLGLPSVINVSDASFQPDGSL
IASSVAPLSQYGDSAGYDPERVFFRCAPEDDVYEMFSTNADDLYSGWYLGGDSAGNSIGLQSAYRTAWPNVLLRLTHVET
GQYFTDVWRERLLGGLDIDSRGFQLVKAKNLSAVRAELFRAPLEFIRYYSPTTASRLYAYTQPAGYIAIKGPGLAYPNVG
ASHNANYLGWHYNWPGAIGLYNDVTLKRYPTCSVTNVTPHVVFPSISLSEINAGANREMPFEVAFKCQTGVINSTASSGT
ALGIRASAGAQAASAALGLRNANGGLSYLVSDRYGQPGMAQGVGIRLLRDGSAMNLLVSEDSAMGSNAETRGWYPVIGNA
SNKTGEAGGISQYSETFRARLEKLTVGSMPSVTPGRVEASAQVVIRVQ

FIGURE 21

RL042

DNA sequence: (SEQ ID NO: 21)

ATGTTCTGCCACGTTGAGGCACGGCGCACCGGCAAACTGCCGCTGGCTCTGGGCGGCCTAGCGCTGGCCTTCGCGGGCCT
GGCCAATGGAGAGGCGCAGTATCGTTTCGACGACAGCCTGTTGATGGGCTCGGGCCTCGCCGGCGGGACCCTAGAACGCT
TCAACCGGGCCAACCAGGTGGACCCCGGAACCTACCATGTCGATGTCTATCTCAACGGCAGCTACGCCAGTCGCACCAGG
ATCGAGTTCCGCCCCCGGGCCGGCGGCGTCAAACCCTGCTTCGGCGAACGCTTCTTGCGCCGGACGCTGGGCGTCCGCCC
CGCCTCTGAGGCCGGCGTGCAAGCGCCTGGAGATTGCCTGGGGCTGGAAGAACGCCTGCCAGGCTCGACCTTCAATCTCG
ACACCGCCCTTCTGCGCCTCGATCTCTCGGTGCCCCAGGCCCTGCTGGATATCAAGCCACGCGGCTACGTGGGTCCCGAC
GAGTGGGACGCTGGCAGTAGCATGGGCTTCGTCAACTACGACGCCAGCTTCTATCGCTCGAGCTTCGACGGAGTAGGCGG
CAACGGCGACTCGGACTATGGCTACCTGGGGCTGAGCGGGGGCATCAATTTCGGCCTGTGGCGCCTGCGCCACCAGTCCA
ACTACAGCTACTCCAGCTATGCGGGAAACACCCGCAGCGACTGGAACAGCATCCGCACCTATGCCCAGCGCGCGGTGCCA
GGCCTGCGCAGCGAACTGACCCTGGGCGAGAGCTTCACCGAGGGCAATCTGTTCGGCAGCCTGGGTTATCGCGGCGTGCG
CCTGGCCAGCGACGACCGCATGCTGGCAGACTCGCAACGCCGCTATGCTCCACAGGTACGCGGTACAGCGAACAGCAACG
CACGGGTGGTCATCAGCCAGAACGGCAAGAAGGTCCACGAATCCGCCGTCGCTCCCGGTCCCTTCGTCATCAACGACCTC
TATGGCACCGCCTACGACGGCGATCTGGATGTCCAAGTGATTGAGGCCGACGGCAGCGTCTCGCGCTTTTCCGTGCCCTT
TTCCGCGGTTCCCGAATCCATGGGCCCGGGCATCTCGCGCTACAGCGCCACCCTCGGCCAAGCGCGCCAGTATGGCGACG
GCAACGACCTGTTCGGCGACTTCACCTATCAGCGCGGCCTGACCAACTCGCTAACCGCCAACCTCGGCTCGCGCCTGGCC
GAGGACTATCTGGCGCTGCTCGGCGGAGGCGTGCTCGCCACGCCCTACGGAGCCTTCGGCTTCAACAGCATCTTTTCCCA
TGCCACGGTGGAGAACGGCCAGCGCAAGCAGGGCTGGCGTGTCGGTCTGAACTACAGCCGGACCTTCCAGCCGACCCAGA
CCACTCTCACCCTGGCTGGCTACCGCTATTCCACCGAGGGCTATCGCGACCTCGGCGACGCGCTTTCGGCGCGCCACGCC
GATGAGCACAACGACTCCTGGAACTCCAGCAGCTACAAGCAACGCAACCAGTTCACCCTGCTGGTCAACCAGGGCCTGGG
GGGCTACGGCAACCTGTATCTGTCCGGAGCCACCAGCGACTACTACGACGGCAAGAGCCGCGACACCCAGTTGCAGTTCG
GCTACAGCAACACTTGGCGCCAGCTCAGCTACAACCTCGCCTATTCGCGCCAGCAGACCACCTGGTACCGCGATCTGAAC
GACGACTACGACCCGTCACTGCCGCCGCAATACAACCTGCGGCACGGCAGCGAACGTAGCAACACCTTAACCCTGACACT
TTCCATGCCGCTGGGGTCCTCCAGCCAGGCCCCGAATCTCAGCGCGATGGCCTCCCGGCGTTCCGGCGACAGCCGCGGCA
GCAGCTACCAGACGGGCCTCAACGGCACCCTCGACGAAGACCGCAGCCTGAGCTACGCGATTGCCGCCGGGCGCGACAGC
GACAACCACGGCAGCGATTTCAACGGCAGTCTGCAGAAACAGACCTCGGTGGCGACGCTGAACGCCGGCTATGCCGAGAA
CAGCAGCTACCGGCAGCTCAACACCGGCCTGCGCGGCGCCGCCGTGCTGCATCGCGGCGGCCTGACCCTCGGCCCCTACG
TCGGCGACACTTTCGCCCTGGTCGAGGCCAAGGGCGCCAGCGGAGCTGGCGTACGCGGTGGTCAGGGCGCGCGCGTCAAC
GGCAATGGCTACGCCGTGGTGCCATCACTCTCGCCCTACCGCTACAACCCGGTCAGCCTCGATCCGCAGGGCATGGGCGA
AGAGGCCGAGCTGCTGGAGACCGAGCGCAAGATCGCGCCATACGCCGGCGCCGCCGTGCATGTGAAGTTCCGCACACTGA
CCGGTCACCCATTGCTAATCCAGGCCCAACTCGCCGACGGCAGCGCGCTACCGCTAGGGGCCAATGTGCTCGACAGCCAG
GGTGTGAACATCGGCATGGTCGGTCAAGGCGGCCAGGTCTATGCCCGCGCCGAGGGCGACAAGGGCCGCCTGCGCGTGCA
ATGGAGCGAACGCCCAGGGGACGCCTGTCTGCTGGATTACGACCTCGACACTGGCCCTCGCCAGGCTATCGAACCCGGAC
AGGCGGTGATCCGCCTGCAGGGCACCTGCACGCCCGTCTCGGAGGCACCATGA

Protein sequence: (SEQ ID NO: 145)

MFCHVEARRTGKLPLALGGLALAFAGLANGEAQYRFDDSLLMGSGLAGGTLERFNRANQVDPGTYHVDVYLNGSYASRTR
IEFRPRAGGVKPCFGERFLRRTLGVRPASEAGVQAPGDCLGLEERLPGSTFNLDTALLRLDLSVPQALLDIKPRGYVGPD
EWDAGSSMGFVNYDASFYRSSFDGVGGNGDSDYGYLGLSGGINFGLWRLRHQSNYSYSSYAGNTRSDWNSIRTYAQRAVP
GLRSELTLGESFTEGNLFGSLGYRGVRLASDDRMLADSQRRYAPQVRGTANSNARVVISQNGKKVHESAVAPGPFVINDL
YGTAYDGDLDVQVIEADGSVSRFSVPFSAVPESMRPGISRYSATLGQARQYGDGNDLFGDFTYQRGLTNSLTANLGSRLA
EDYLALLGGGVLATPYGAFGFNSIFSHATVENGQRKQGWRVGLNYSRTFQPTQTTLTLAGYRYSTEGYRDLGDALSARHA
DEHNDSWNSSSYKQRNQFTLLVNQGLGGYGNLYLSGATSDYYDGKSRDTQLQFGYSNTWRQLSYNLAYSRQQTTWYRDLN
DDYDPSLPPQYNLRHGSERSNTLTLTLSMPLGSSSQAPNLSAMASRRSGDSRGSSYQTGLNGTLDEDRSLSYAIAAGRDS
DNHGSDFNGSLQKQTSVATLNAGYAENSSYRQLNTGLRGAAVLHRGGLTLGPYVGDTFALVEAKGASGAGVRGGQGARVN
GNGYAVVPSLSPYRYNPVSLDPQGMGEEAELLETERKIAPYAGAAVHVKFRTLTGHPLLIQAQLADGSALPLGANVLDSQ
GVNIGMVGQGGQVYARAEGDKGRLRVQWSERPGDACLLDYDLDTGPRQAIEPGQAVIRLQGTCTPVSEAP

FIGURE 22

RL043
DNA sequence: (SEQ ID NO: 22)

ATGAATACTTTTCCACTGCCTCCGCTCCGTGCGGCTACGCTGGCGCTCGCCCTGCTGATACCCGCCATCCCGGCTCAAAG
CAGCGTGGTGATCATCGGTACTCGCGTGATTTATCCCGGCGACGCCCGGGAAAAGACCGTGCAGATGATCAATCAGGACG
CATTCCCCAACGTGATCCAAGCCTGGATCGACAACGACGACCCCTCCTCCACCCCGGAGACTGCAAACGCGCCCTTTCTG
GTCAGCCCAGCGGTGACGCGCATAGCCCCCGGCAGCGGCCAGACCCTGCGCCTCCTGTATACCGGGCTCCCGCTGCCCGA
GGATCGCGAATCGTTGTTCCATCTCAATGTGCTGCAGATCCCGCCCCGCGACCTGGCCAAGGCCGAGCGCAACCAGATGC
TGCTGATGCAGCGCAGTCGACTGAAGCTGTTCTATCGCCCCGCCGCGCTGCTTGGCGGCTCGGAGCAGCTAGTCGAGCAG
TTGCACTTCAGCCTGGTGCAGGCGAGCGGCAACTGGCGTGTGCGGGTGGACAACCCCAGCGGCTACTACGCCTCCTTCGC
CGGCGCGATGCTGAGCATCGGCGAACGTCGCTGGCGGCTGCTGTCGAGCATGGTCCCGCCCAAAGGCCAGGCCGAGTGGG
CGGCGGAACGCCCTTCGCCGCTCGCCCCAGGACCGGTCCAGTTGAACGCCCTCTTGATCAATGACTACGGCGCGCGAATG
GAGGTCCAGCATGTTCTGCCACGTTGA

Protein sequence: (SEQ ID NO: 146)

MNTFPLPPLRAATLALALLIPAIPAQSSVVIIGTRVIYPGDAREKTVQMINQDAFPNVIQAWIDNDDPSSTPETANAPFL
VSPAVTRIAPGSGQTLRLLYTGLPLPEDRESLFHLNVLQIPPRDLAKAERNQMLLMQRSRLKLFYRPAALLGGSEQLVEQ
LHFSLVQASGNWRVRVDNPSGYYASFAGAMLSIGERRWRLLSSMVPPKGQAEWAAERPSPLAPGPVQLNALLINDYGARM
EVQHVLPR.

FIGURE 23

RL044

DNA sequence: (SEQ ID NO: 23)

ATGAAACCTCAAAGTACTGCCCTGACTATCGCCGCATTTCTCGCATTGCCGGGTATCGCGGCGGCTGCCAATACCATCAC
CTTCCACGGAGAAGTGACCGACCAGACCTGTTCCGCCGTCGTCGACGGACGAACCGACCCGACCGTGATACTCGACACCG
TACCGGTAAGCGCTCTTGACGGCGCAGTCGGCAAACCCGCCGGGGAAACCAGCTTCACCCTGCAACTGACCGGTTGCGCC
GCTCCGGCGGCCGATGCCGAGGAGCACTTCAGCGTGATGTTCCAGGCGGTCAATCCGACCAGCGCCGGCAATCTGACCAA
TACCGCGTCCGCCGGCGCCACCGGCGTAGCGCTGCAGCTACTGACGGCACCGGGCGGCAGCGAGGTCAATCTGGCCGGCG
GGTCGGCCGTGGCTGCCGGTGACATCGTGCTCGCAGGAGGCGAGACCAGCACCAGCTACGACTATGCCGTCCGCTACATC
TCCGAAGCGACCACCGTCACTCCGGGACCGGTGCTCGGCTCGGTGACCTACACCCTGCGTTACGAGTAA

Protein sequence: (SEQ ID NO: 147)

MKPQSTALTIAAFLALPGIAAAANTITFHGEVTDQTCSAVVDGRTDPTVILDTVPVSALDGAVGKPAGETSFTLQLTGCA
APAADAEEHFSVMFQAVNPTSAGNLTNTASAGATGVALQLLTAPGGSEVNLAGGSAVAAGDIVLAGGETSTSYDYAVRYI
SEATTVTPGPVLGSVTYTLRYE

FIGURE 24

RL045
DNA sequence: (SEQ ID NO: 24)

AGTCCGCACGGTAGTGACGACTGGAAGCGCTTCTGTGCTGCCAACAACCTGGAGCCCAGCATGAGCCGGCGCGGCAATTG
TTGGGATATGCCGTGGCGGAATCCTTCTTCAGTAGTTTGAAGAAAGAGCGTATCCGCAAACGCATCTACAAAACCCGAGA
CATGGCCCGGGCGGATGTTTTTGACTACATCGAGGTCTTCTACACCCGAACCCGGCGGCACAGTCATCTGGGTGGCGTCA
GTCCCGAGGCCTTTGAAAGCGCCTCG

FIGURE 25

RL046
DNA sequence: (SEQ ID NO: 25)

ATGGCTGAAGTCACTCAACGAGCAGAGCAGCAACAAGAGAGCCAGAAGACCCTTCTCGGCACCATCATCAGTACGCCCTT
CCAATTTCTCGGCGTGATGTTCGGGTCGCTGATCGGCGCAATCATCGTGGAGTGGGTTTGCCTGTATTTCTTCTGGCCTG
ACGCGGGCTGGAAGCATGCCCAGGCCATGTTTGAGTACGAACTCAGTTGGCTGTCGCAGGGGCTGCTACACAGCGTCGTC
GTGCAGGAGCCAGGTCGAACCGCCACCTGGCTGGCCCAGTTGGCCTATGACTGGTTGTTCGTGAAGACCGGGATGGTCGA
CTGGATGACCAACATGACTACCATCGCGCAGGCCGGGCCACGGAGCCCGCTGGACGTTCGCTATCTCACCGCCCAGGGTG
TCTCCACGCTGCAGAACTACGGCCTGGCCGCGCTGTACACGGTGCTGACATTCGTCGTGCGCCTGGTGATCCTGGTCATG
ACGATCCCGTTATTCGTGATGGCCGCGTTCACCGGCCTGGTGGACGGCCTGGTGCGCCGGGACCTGCGCAAGTTCGGCGC
CGGCCGGGAGTCCAGCTACCTCTACCACAAGGCGCGCGGCAGCATCATTCCGCTAGCGGTCGTCCCTTGGACGCTCTACC
TGGCAATTCCCATCAACATCAATCCCCTGCTCATCCTGTTGCCCTGCGCCGCACTGCTCGGCGTAGCGGTGTGCATCACA
GCATCCACCTTCAAAAAGTACCTATAG

Protein sequence: (SEQ ID NO: 148)

MAEVTQRAEQQQESQKTLLGTIISTPFQFLGVMFGSLIGAIIVEWVCLYFFWPDAGWKHAQAMFEYELSWLSQGLLHSVV
VQEPGRTATWLAQLAYDWLFVKTGMVDWMTNMTTIAQAGPRSPLDVRYLTAQGVSTLQNYGLAALYTVLTFVVRLVILVM
TIPLFVMAAFTGLVDGLVRRDLRKFGAGRESSYLYHKARGSIIPLAVVPWTLYLAIPININPLLILLPCAALLGVAVCIT
ASTFKKYL.

FIGURE 26

RL047
DNA sequence: (SEQ ID NO: 26)

ATGGCTGGCCAGTACCCGTTGGAAGCGCTCTTGCGGCCTGCCGTGGAGCTCTACACCACCACCGTGTGCTTCACCGCAGC
CGCGCTCTGCATCGTCGCGCCGTGGACGTTCTCCCTCACTCCGCTGTTCGGCATCGTGGCCGCGCTGTGCTTCGCCTGGC
TGGGTATCGTGCGGCTGAAGCAGGCCGGCGTGGTGCTCCGCTACCGGCGGAACATTCGCCGACTGCCGAAGTACACGATG
ACCAGCGCCGAGATGCCGGTCAGCAACGAACACCTGTTCATCGGTAAAGGATTTCGCTGGACGCAGAAGCATACGCAGCG
CCTGGCAGATACCTACCTGCCCCAGTTCGCCTCTTACGTCGAGCCCTCGCCCCTCTACGAGCGCGCGCGCCGGTTGGAGA
AGCAGCTCGAGTTCGCCCCCTTCCCCCTGAAGCTGGTCGCCAAAGCCACTGCCTGGGACGTGGCCTGGAACCCCGCACGG
CCGCTGCCGCCCGTGGGCGGTTTGCCTCGGCTCCATGGCATCGAGCCGCGCGAACAGGACGTAGGCCTGCAACTGGGCGA
GCGCGTCGGCCACACACTGGTACTCGGCACCACGCGGGTGGGTAAGACGCGCCTCGCCGAGCTGTTCATCACCCAGGATA
TTCGCCGCACTCACTGCCGGGTACGACGCCGGCGGGTGAAGATGGGCCGGCGGACCCAGACGGTTCACCACGGCTATCGG
CGCCGGCGCGCAGAGGAGCAGCCGGACTACGAGGTGGTGATCGTCTTCGACCCGAAAGGCGACGCTGACCTGCTGAAGCG
TATGTACGTGGAATGCGAACGTGCCGGCCGCCTGGACGAGTTCTACGTGTTCCACCTCGGTCATCCTGACCTGTCGGCAC
GCTACAACGCCGTCGGCCGGTTCGGTCGGATCTCCGAGGTCGCCACCCGCGTCGCCGGCCAGCTCTCCGGCGAGGGCAAC
AGCGCGGCGTTCCGCGAGTTCGCCTGGCGGTTCGTCAACATCATCGCCCGCGCGCTGCACGCGCTGGGTATCCGGCCTGA
CTACCAGCAGATCCTCCGGCACGTCGTGAACATCGATGCGTTGTTCGTCGAATATGCGCAGAAATACATCAGCGAGCACG
ATCCCAGGGCCTGGGACACCATCATCCAGATCGAGGGCAAGCTCAACGACAAGAACATCCCGTTCAACATGAAAGGACGG
CCCCTGCGGGTCGTAGCCATCGACCAGTACCTGACACAGAAACGCATCGCCGACCCGGTCATGGAAGGCTTGAAGAGCGC
CGTGCGCTACGACAAGACCTACTTCGACAAGATCGTGGCCTCGCTGCTGCCGCTACTGGAGAAACTCACTACCGGGCGGA
TCTCGGAGCTTCTTTCGCCCAACTACGCGGACCTCAACGATCCGCGGCCGATCTTCGACTGGATGCAGGTCATCCGCAAA
CGCGCCGTGGTCTACGTCGGCCTCGACGCACTATCGGATACCGAGGTCGCCGCCGCGGTGGGCAACTCCATGTTCAGCGA
CCTGGTCTCGGTAGCGGGTCACATCTACAAGCATGGTGTCGATGACGGCCTGCCCGGCTCGCTCGCCAGCGGCAAGGTCC
GCATCAACCTGCATGCCGACGAGTTCAACGAGCTGATTGGCGACGAGTTCATCCCCATGGTCAACAAAGCGGGCGGCGCC
GGCGTGCAGGTGACGGCCTACACCCAGACCATGAGCGACATCGAGGCCAAGATCGGCTCCCGCGCGAAGGCCGGTCAGAT
CATCGGCAACTTCAACAACCTGTTCATGCTGCGGGTGCGCGAGACCGCCACGGCCGAACTCCTTACCAATCAGCTCCCCA
AGGTCCAGATCTACACCAGCACGCCGGCGAGCGGCGCCAACGACGCGATCAACAACAACAAGAAGGTAGCCTTCACCTCC
AGCTCGCACGACCAGGTGCAGATGACCAGCGTGCCGATGCTCGAGCCGGCCCACATCATTGGTCTGCCCAAAGGACAAGC
GTTCGCGCTACTCGAGGGCGGCAATCTCTGGAAGATCCGAATGCCGCTGCCGGCGGTCGCCCCCGACGAGGTGATGCCGA
AAAGCCTGCAGGAGCTGGCTGCCGGTATGCGCAAGGGCCAGGCCGCCAACAGCGAGTGGTGGGAGGCGCCGGGATACTCC
GCCCTGCAGGATGGTCTGCCCCAGGACCTGGTCGACGATTTCCGTCACCTCGGCACCGGTGAGGATGCCGCCTGA

Protein sequence: (SEQ ID NO: 149)

MAGQYPLEALLRPAVELYTTTVCFTAAALCIVAPWTFSLTPLFGIVAALCFAWLGIVRLKQAGVVLRYRRNIRRLPKYTM
TSAEMPVSNEHLFIGKGFRWTQKHTQRLADTYLPQFASYVEPSPLYERARRLEKQLEFAPFPLKLVAKATAWDVAWNPAR
PLPPVGGLPRLHGIEPREQDVGLQLGERVGHTLVLGTTRVGKTRLAELFITQDIRRTHCRVRRRRVKMGRRTQTVHHGYR
RRRAEEQPDYEVVIVFDPKGDADLLKRMYVECERAGRLDEFYVFHLGHPDLSARYNAVGRFGRISEVATRVAGQLSGEGN
SAAFREFAWRFVNIIARALHALGIRPDYQQILRHVVNIDALFVEYAQKYISEHDPRAWDTIIQIEGKLNDKNIPFNMKGR
PLRVVAIDQYLTQKRIADPVMEGLKSAVRYDKTYFDKIVASLLPLLEKLTTGRISELLSPNYADLNDPRPIFDWMQVIRK
RAVVYVGLDALSDTEVAAAVGNSMFSDLVSVAGHIYKHGVDDGLPGSLASGKVRINLHADEFNELIGDEFIPMVNKAGGA
GVQVTAYTQTMSDIEAKIGSRAKAGQIIGNFNNLFMLRVRETATAELLTNQLPKVQIYTSTPASGANDAINNNKKVAFTS
SSHDQVQMTSVPMLEPAHIIGLPKGQAFALLEGGNLWKIRMPLPAVAPDEVMPKSLQELAAGMRKGQAANSEWWEAPGYS
ALQDGLPQDLVDDFRHLGTGEDAA

FIGURE 27

RL048

DNA sequence: (SEQ ID NO: 27)

ATGACTACTCATCTGATCACCCTAGTCATCAAGCAGCCGAGCGACGCTCAGGCACGCCAACTCATGTACCAGGAGTTGCT
CGGACTGATCTCACGCTACGGCGGTGAGGTGACGTCCAAGGCCTTGGAGGACGAGTCGACCCTCTGCGAGCTGCTGGTGC
AGATGCTGCCTGATCATGAGGTAGAGCAAGCCAGGAAACAGGTGCTCGAACTTCATGCCAAGGGCCGCCTGCAGGCGCCG
GCAAGCCTGAAGGTGTAA

Protein sequence: (SEQ ID NO: 150)

MTTHLITLVIKQPSDAQARQLMYQELLGLISRYGGEVTSKALEDESTLCELLVQMLPDHEVEQARKQVLELHAKGRLQAP
ASLKV

FIGURE 28

RL049

DNA sequence: (SEQ ID NO: 28)

ATGAAGAAGTTCCTTGCCACGCTGGCATTTTGCACGGCGTTCGCGACTCAAGCCTGGGCCGCCGGGCTGATCGTTGTCGA
AGACCTCGGCGGCGCCTCGGCGCTCCCCTACTACCAGGGCCTGGATCCGCAGCCATCCGCTTCCGCACCAGGACCTGGCG
ACCTGGGCGTCCGTGGCTCAGGTGCGTTTCCAGTTCGCTCCGCCCGCCTATCGCCAGGACGGGTCCAGGGGCGCGCCATC
AACGCTCCAGGCCTGCAACTGCTGTTCCTGGTCGGCGACGACACGCTGTCTCGAACCTGGCTGAAAGAGCGAGGCGACGA
GCTTCGAGACCTCCAAGCCGTGGGCCTGGCAGTGAACGTGGCCAGCGAAGCGCGCCTGACGGAAATCCGGGCCTGGGGGA
AAGGACTTCAGATATTGCCGGCGCCGGCGGACGACCTGGTCGACCGGCTAGGGCTGCAGCATTACCCCGCCCTCATCACA
TCCACCGCCATCCAGCAGTAG

Protein sequence: (SEQ ID NO: 151)

MKKFLATLAFCTAFATQAWAAGLIVVEDLGGASALPYYQGLDPQPSASAPGPGDLGVRGSGAFPVRSARLSPGRVQGRAI
NAPGLQLLFLVGDDTLSRTWLKERGDELRDLQAVGLAVNVASEARLTEIRAWGKGLQILPAPADDLVDRLGLQHYPALIT
STAIQQ.

FIGURE 29

RL050
DNA sequence (SEQ ID NO: 29)

ATGGCAACGTCTGTAGTTCGAGCCCTCCAGTTGGCCACCCTGCTGGTCCTGGTCAACATCGCTCAGGCCGCCGTGGATCC
ACCGCCGGCGTACAAGCAAATCGCCCTGCCCAAAGGGGTTCCGGCCGAGGTGCTCTACTCGGTCGCGCTGACCGAGAGCA
AGGTCCTGCTGCGCGGCGAATACGTTCCCTGGCCCTGGACATTGAACGTCGCCGGGAAATCTTACTACTACGCGACCCGC
ACCGCCGCCTGCACAGCGCTACTCGCGGCGATCAACCTCTACGGGGCCAAGAGCGTCGATTCCGGCCTCGGCCAGGTCAA
CATCGGCTGGAACGGACATCGTTTCTCCAGCCCCTGCGAGTCCCTGGATCCGTACAAGAACCTGGACGCCACCTCCGACA
TCCTGATCGAGCAGCGGGACGCCCTGTATGCATCCGCCCCGGGAAGACCGGTGGACTGGATCCAAGTTGCCGGCCGCTAC
CACCGCCCCGCCGGCGGCGCGCCTGCCGCCAAATACCGTAGGACGGTTTCCCGCCACCTTAGCCAAGTTCTCGGCGTCAA
CCTACTGGTGACCAATCCATGA

Protein sequence (SEQ ID NO: 152)

MATSVVRALQLATLLVLVNIAQAAVDPPPAYKQIALPKGVPAEVLYSVALTESKVLLRGEYVPWPWTLNVAGKSYYYATR
TAACTALLAAINLYGAKSVDSGLGQVNIGWNGHRFSSPCESLDPYKNLDATSDILIEQRDALYASAPGRPVDWIQVAGRY
HRPAGGAPAAKYRRTVSRHLSQVLGVNLLVTNP

FIGURE 30A

RL051
DNA sequence: (SEQ ID NO: 30)

ATGATCAGAACCGTATCGCTCCTGTCCGGCCTGATGCTGCTGCTGAGCTATCCCGCAGCCGGCCAGGAGGCGGCGGCAAG
CCGAGAGGCCAGCAGCCAACTGTCCGGTAGCCAACTCGGCACGCTGAAACAACAGACATCTCAGAGCGACCTGGCCCAGG
AGTGGGGACTGAACCAACAGGAATGGACCCGCTACCAGACGCTCATGCAAGGCCCCCGGGGCGCTTACTCGCCTGGTATT
GATCCGCTGACCGCGCTGGGCATCGAGGCGCGATCGGCAGAGGAACGGCGGCGGTATGCCGATCTACAGGTCCAGGCCGA
ACGGCGCCGGGTCGAGAAGGAACTCGCCTACCAGCGCGCATACGACGAAGCCTTCGCCCGCGCCTATCCAGGCGAGGGGG
TGATCCGCCTCACCGAAAGCAGCACAGCCAACCCGTCGGGCACGCCGAACATGAGCCCAGCGTTGCAGAGCAGCGGGCGC
CTGGCCCTGTTCGTCCAGGACAACTGCACCGCCTGCATCCAGCGGGTCCGCGACCTGCAACATGCAGAAAAGGAGTTCGA
CCTCTACTTCGTCGGTAGCCAGAACGACGCAGAGCGAGTGCGGCGCTGGGCAATCCTCGCCGGCATCGACCCGAAGAAGG
TTCGCAGCAAGCAGATCACGCTCAATCATGACGAGGGCCGCTGGATGGCCCTAGGACTGGGCGGAGCCCTTCCCGCCCTG
GTACAGGAGGTGAACGGCCGATGGCAACGTCTGTAG

Protein sequence: (SEQ ID NO: 153)
MIRTVSLLSGLMLLLSYPAAGQEAAASREASSQLSGSQLGTLKQQTSQSDLAQEWGLNQQEWTRYQTLMQGPRGAYSPGI
DPLTALGIEARSAEERRRYADLQVQAERRRVEKELAYQRAYDEAFARAYPGEGVIRLTESSTANPSGTPNMSPALQSSGR
LALFVQDNCTACIQRVRDLQHAEKEFDLYFVGSQNDAERVRRWAILAGIDPKKVRSKQITLNHDEGRWMALGLGGALPAL
VQEVNGRWQRL.

RL052
DNA sequence: (SEQ ID NO: 31)
ATGAAACGCCCATCCCCTGCATCAATGATTCTTGGCCTCTGTTTGACGGCAATGGCCGGCCTGCTGAGCTACCAGCAGTA
CCAACTCGTTCAGCTCCGATCAGGCGTGGACAGTGCCGCGGAAAAGGCCTCGCTGGAGGCGATCCTGGCTCGCTTGAGTC
GAGTCGACGAGCGCCTCGACGCCGTGGATGGACAGCACCTGGTCAGCAACGAGGACTTCCGTTCAGGCCAGCAGGCGCTG
TCCAACCGAATTGACGCTGCGCAGGCGTTCGCCAAGCAGGCCTCCGATGCCGTCGAGAACCTGGCTCAGACCACCGCCTC
GGCCGGCGACCTCTTGGTGCTCAAGGCAACCGTGGAGACACTGGACGGTTCTGTCCGCACGCTTCAAGAAAAGCAGGCCA
AGGCGCCGCCGCTGATCGTGCCAGCGCCAAAACGCCCCATACCCGCCAAGCCCAAGCCGAAACCCAAACCGATGGAGCCC
CCGCCCTTCTCGATCCTTGGCGTGGAGTATCGCGGGGGAGAACGGTTTCTGTCGGTTGCACCTCCGGGATCCACCCAGCT
CAGCCAGATCTACCTCATTCGCCGGGGAGATGCCGTCGCCGGCACGACCTGGCGACTGACCGACCTTGACGATGGTACCG
CGCACTTCGACGTCGCCGGCACCTCGCGCAGCGTTCGCATCCAACCATAG

Protein sequence: (SEQ ID NO: 154)
MKRPSPASMILGLCLTAMAGLLSYQQYQLVQLRSGVDSAAEKASLEAILARLSRVDERLDAVDGQHLVSNEDFRSGQQAL
SNRIDAAQAFAKQASDAVENLAQTTASAGDLLVLKATVETLDGSVRTLQEKQAKAPPLIVPAPKRPIPAKPKPKPKPMEP
PPFSILGVEYRGGERFLSVAPPGSTQLSQIYLIRRGDAVAGTTWRLTDLDDGTAHFDVAGTSRSVRIQP.

RL053
DNA sequence: (SEQ ID NO: 32)

ATGCCGCGCCGCTTGATCCTCTCGGTCACGGAGCGGGATATCCTATTTGCACTGCCGGTAAGCCGAATGACCTCACTGAC
TACTCCACCTCAACGAGTCCGCCCCATCGTCGATCCGCCAGCGACGCGGCGATGCCAATCACTTGGTTTTTTCGGTGCAG
GTCAGCCTGCTGTGCTATCCAGCGTTCAGCCCTGATGCGCGACGAAGAGCCCCCGAG

RL054
DNA sequence: (SEQ ID NO: 33)

ATGGCCGAAGCTATCAGAAAGGATGCAATGATGACAAAACTCTACTTTGATCTTCTGAACTCGCCTGCCGAGGCTCATTC
GTCGATACAAAAGTCTTTATCTGTGCAGGCAATCTCCACAACTGTCCCAATACTGGAGTTTCCTTCGGAAACCGTATACG
CCTATGCATCGTACATAAATGCATTAAGTATCGGTCAACGCATAGATCCTGCATTCACCCAGAGCTTAACGAGTGCCATA
TCCAACCTGGCAGGTCGCCCGATTGCAGTAAGCGACATTTACCAAAAAATTCATGAAACCACACTGAGAACACCTGTTGA
GATGGGCGTTCGTCCTAATAGCATCACCTTTGAGGAGTATCAGGCCACCATAAATCAGCAAGCCATCAACATGGTTCAAG
ATATGCAGGATGGAGACAAAGGTGAGAAGGTGGAGGCCCTCCAGGCCAATATGCAGTTCCTGTATGGACAGGAGATAAAT
ACTGATTTCATCGCTCGTAATGAACTCGCTGCTGGGCAGAGAGCGAAAACCGTCGCAATAGTTCAGGGGCATATCACCAT
CGGGTACGGCTTCGATACCTTCGTGCATGAAGCGTCCGAGCTAAACTCTTTGAATCTTGTTGGTTCTACGCGACAGAAGG
TATTACCTGCATTGCAGCTATCAACGTCCGACCCAGGCTTCTGGAGCGTCTATGCCTTGCTGGGACAAAGTCTCACGGAT
GACGATGGGCTATTACTCTTTAGTGCCAAAGCGCGAGCTGTTGTTCAACGCATAGCAAGCAACCAGTTTGCAGGTAAGTG
GAATGGGCTACCCCCAGCTATCAAAACGGTTGCGCTTGATCTATATTATCAATATGGGCAGACTGGTAATTTTCCAAAAT
TTCAACAAGCTATAAATAGCCATGATTGGCCGGCAGTCATCCATGAACTTAGAAACTGGAATGGTGTACCGAATGATCCT
CTCCAGTTCATTACAAAACGATTGGAAGAGCGAGCCAAGTATCTGGCAATATCCTTCAACTATGAGCAATGA

Protein sequence: (SEQ ID NO: 155)
MAEAIRKDAMMTKLYFDLLNSPAEAHSSIQKSLSVQAISTTVPILEFPSETVYAYASYINALSIGQRIDPAFTQSLTSAI
SNLAGRPIAVSDIYQKIHETTLRTPVEMGVRPNSITFEEYQATINQQAINMVQDMQDGDKGEKVEALQANMQFLYGQEIN
TDFIARNELAAGQRAKTVAIVQGHITIGYGFDTFVHEASELNSLNLVGSTRQKVLPALQLSTSDPGFWSVYALLGQSLTD
DDGLLLFSAKARAVVQRIASNQFAGKWNGLPPAIKTVALDLYYQYGQTGNFPKFQQAINSHDWPAVIHELRNWNGVPNDP
LQFITKRLEERAKYLAISFNYEQ.

FIGURE 30B

RL055
DNA sequence: (SEQ ID NO: 34)
ATGAACAACACAGTGAGCGAAACGCAACAGATCAATATTTACCAAAATCCGGGGCAGTCTATTTCCGGTCTCTACAAGGG
GCTGGCTAACCAGTGCTCTCCTGGCCAGCCATTTCCAGAGGTACAGCTTGTGGAGGCTTGGGATATCCCTCTCGTACTCC
ATCCGGAGTTTGTGCCTAACGGAGATGTCTCGAAAATCGATAAGGAGTACGGAACGATCCTTGCTGCTGAGTCAGCTCAG
GTTATCCTGCTTCAACTCCAAATGGCTCAAGACAAGGCTAAGGCGTGCGGGGAGGTTACAGCCTTGATCAGTTCTGTCTC
CTCCAATCTCAATACCATTAAGAGTCGTCATGGTGCTAATTATCTAAACCTGCTGAAACAATCACCGAACCGATACCCGA
CTAGCGTCGGAGTTGAGATCATGTCAGGTGGCAGTCCGAACCAGGATTCTGGAATCGAGGTCTCTTACGGTGCCAGTCTC
GGCCGTCTAACTCAATCACAACTTCAGGCGATGAATCTGCCTGCCAGTCTCAAACAGTTGCTCACTCAGGGAATCGGTGT
GAAGCTTTCTCAGCCTGAATATTGGCCTGCTTACAACAACATAGCCACTGGTATTCGTTATACAACCGGAGTGGCGATAA
CGTTGGCCTATTGGGCCACGGTTTAG

Protein sequence: (SEQ ID NO: 156)
MNNTVSETQQINIYQNPGQSISGLYKGLANQCSPGQPFPEVQLVEAWDIPLVLHPEFVPNGDVSKIDKEYGTILAAESAQ
VILLQLQMAQDKAKACGEVTALISSVSSNLNTIKSRHGANYLNLLKQSPNRYPTSVGVEIMSGGSPNQDSGIEVSYGASL
GRLTQSQLQAMNLPASLKQLLTQGIGVKLSQPEYWPAYNNIATGIRYTTGVAITLAYWATV

RL056
DNA sequence: (SEQ ID NO: 35)

ATGACCCAAGCTGCGAAAATACCAGCAAATGAGTACTCATTGGGGGATGGAAGAGGCTACATCAATATCTGGCCGGAAAA
GGATGAGGCTCAGGCATTTCTTATCCATAATGATGGGCCTAATGGGGCTACATGCAGCCTTAAAGGCACTCTTAGAGATA
ATAAAGGAGTGGTGCATTCGCCGTATTCCTCTGCTTCATGTTTGCTAAGTATCACCCAGACAGGGCTGCTGTCAGTAAGC
GTCAAACGTGAGGAAAATTCGCCAAGCTGCTCTGCATGGTGCGGTCCTAGAGTTTGGTTTGAAGGAGCCTATAGCGTCCC
GCCCAAGGGCTGCTACTATATGCAAATAAGGAAAAAAACTCGACAAATGTTGGGTATGATTGAGAAAAAAGAGCTTGATG
CCGCTCGCGCCTTATCAAATAAGCTTTTGTCAGACTGCGCAACCGAGCTAGCCTATCCTGCCAAGATATACTTGACGAAC
ACACTTGCCATGATCAGTGCTGAAAAGGGAGAGAATGCTCGCTGTTTGGAGTATGCCCATCGGGTGCAAAAGCAAATTCC
TGTAAGAGATGACGGCCAACCGGCTGAAGACTTGCTCCCGGCGGAGCACGCTTTCGCTATGGAACAACGCGCCAAGGCTG
ATGCTCTGTCTGAGCGATGCAGCGACGAGAAATAA

Protein sequence: (SEQ ID NO: 157)
MTQAAKIPANEYSLGDGRGYINIWPEKDEAQAFLIHNDGPNGATCSLKGTLRDNKGVVHSPYSSASCLLSITQTGLLSVS
VKREENSPSCSAWCGPRVWFEGAYSVPPKGCYYMQIRKKTRQMLGMIEKKELDAARALSNKLLSDCATELAYPAKIYLTN
TLAMISAEKGENARCLEYAHRVQKQIPVRDDGQPAEDLLPAEHAFAMEQRAKADALSERCSDEK

RL057
DNA sequence: (SEQ ID NO: 36)

GTGCTGGTAGAGCGTTTGCCGACTGATGTTGAATTCGCGGGCGAGCTGAGCCTTGGGCTCGCCGGCCGCTGCCCGCAGCC
CCAGGGTAGCACCTGCTTGTCGGACAAGGCCTCTTTGCGGCCCCGGTACGCGCAGAGCTTGATATCCTCGCGCTACCGCG
CTGGTGCTGCTTGTATGCTGCTCAGCAAGCCAGCCGCCGGTCTTTTCAGGGTAAGCGTACGGCCAATACACCTTTACTTA
GGTTGA

Protein sequence: (SEQ ID NO: 158)
VLVERLPTDVEFAGELSLGLAGRCPQPQGSTCLSDKASLRPRYAQSLISSRYRAGAACMLLSKPAAGLFRVSVRPIHLYL
G

RL058
DNA sequence: (SEQ ID NO: 37)

ATGGATATTCGCCTGGAGATTTTAGCGCTTGAACAGCTGTTGCTAGAGCCGGAATCGAGAAAGAATGATCGACTGCTTAA
ACAGCTGCTTACCGAAGACTTCGTTGAATTTGGAGCTATCGGCAAAAGCTGGACGAAAGCGGAGGTGATCGTGGGACTAA
AATCCCAGACTTGGATCAAAAGGACAATCGAGGATTTCAAACTGCGTGTGCTTGCAGATGGTGTCGCGTTAGCAACGTAC
CGATGCCGTCATCAAAATGCTAATGGCGATGAGTCGTTATCAATGCGTAGCTCTGTTTGGAAAACCTACGAAGATGGTTG
GCACATGGTGTTTCACCAAGGCACGAGGGTCTCCGAGTAG

Protein sequence: (SEQ ID NO: 159)
MDIRLEILALEQLLLEPESRKNDRLLKQLLTEDFVEFGAIGKSWTKAEVIVGLKSQTWIKRTIEDFKLRVLADGVALATY
RCRHQNANGDESLSMRSSVWKTYEDGWHMVFHQGTRVSE.

FIGURE 30C

RL059
DNA sequence: (SEQ ID NO: 38)
ATGACTTCCTCGCCCAACCTTGACCAGATGACCCCGGAACAGCTTCGTGCCTTGGCGGCACAGGCGTTGCAGTTGCAATC
CCAGGTCGAGGCGATGAGCAGGAAAATCCGCAACAATGAAACCCTCATCGAACAGTTCAAGTTCGAAATCGCTCTGCTCA
AACGCCACAAGTTTGCCAAGCGCAGCGAGCAAATCAGTTCGGCGCAAGGCAGCTTGCTGGATGACCTGCTCGACACCGAC
CTTGAAGCTATCGAGGCCGAGCTGAAACAACTCCTTCCAGCTTCGCCACAAGCCGAGCCACGGCAATCCCCGAAACGTTC
GCCATTGCCGCCGCAGTTCCCGCGCACGGTGATTCGCCACGAACCTGAAAATACCCAATGCGCCTGCGGCTGCCAACTTC
AACGCATCGGCGAAGACGTCAGCGAGAAGCTGGATTACACGCCGGGCGTGTTTACCGTCGAGCAACATGTGAGGGGCAAA
TGGGCCTGCCGTCAGTGCGAAACCCTGATCCAGGCGCCGGTGCCAGCCCAGGTTATTGATAAAGGCATCCCGACCGCAGG
TTTGTTGGCCCACGTGATGGTGGCCAAGTTTGCCGATCACTTGCCGCTGTACAGACAGGAAAAAATCTTTGGCCGCGCCG
GGCTGCCAATTGCCCGCTCGACCCTGGCGCAGTGGGTCGGACAAACTGGCGTGCGGCTTCAGCCACTGGTCGATGCACTG
CGTGAAGCCGTGCTGAACCAGGACGTGATCCACGCCGATGAAACACCGGTGCAAATGCTTGCACCAGGCGAGAAGAAAAC
CCACCGGGTCTATGTCTGGGCCTACAGCACGACGCCGTTTTCGGCGCTCAAAGCGGTGGTTTACGACTTCAGCCCAAGCC
GTGCCGGAGAACATGCACGCAACTTCCTAGGCGACTGGAATGGCAAGCTGGTCTGCGACGACTTCGCTGGATACAAGGCC
GGTTTTGAACAAGGCATCACTGAAATCGGCTGCATGGCTCATGCTCGCCGCAAGTTCTTCGACCTGCATGTCGCTAACAA
AAGCCAACTGGCCGAACAGGCGCTGCACTCAATTGGCGGTTTGTACGAGGTTGAACGCCAGGCTCGGGACATGAGCAACG
AAGACCGTTGGCGAATACGTCAGGAAATGGCGGTACCGATCAGCAAAACACTGCATGACTGGATGTTGGCCCAGCGCGAC
CTGGTGCCCAACGGCTCGGCCACAGCTAAAGCCCTCGACTACAGCCTGAAACGCTGGGGAGCGCTGACGCGCTACCTGGA
CGATGGGGCTGTGCCCATCGACAACAATCAGGTGGAGAACCAGATACGGCCGTGGGCGCTCGGACGCTCGAACTGGTTAT
TTGCCGGATCGCTGCGCAGTGGCAAACGAGCAGCAGCTATCATGAGCCTGATCCAGTCCGCTCGCATGAACGGGCATGAT
CCGTATGCCTACCTGAAGGACGTGCTAACTCGCCTGCCGACGTTACGGTCGAAAGACATCAGCCAGTTGCTGCCGCATCA
GTGGGTACAGATCTAG

Protein sequence: (SEQ ID NO: 160)
MTSSPNLDQMTPEQLRALAAQALQLQSQVEAMSRKIRNNETLIEQFKFEIALLKRHKFAKRSEQISSAQGSLLDDLLDTD
LEAIEAELKQLLPASPQAEPRQSPKRSPLPPQFPRTVIRHEPENTQCACGCQLQRIGEDVSEKLDYTPGVFTVEQHVRGK
WACRQCETLIQAPVPAQVIDKGIPTAGLLAHVMVAKFADHLPLYRQEKIFGRAGLPIARSTLAQWVGQTGVRLQPLVDAL
REAVLNQDVIHADETPVQMLAPGEKKTHRVYVWAYSTTPFSALKAVVYDFSPSRAGEHARNFLGDWNGKLVCDDFAGYKA
GFEQGITEIGCMAHARRKFFDLHVANKSQLAEQALHSIGGLYEVERQARDMSNEDRWRIRQEMAVPISKTLHDWMLAQRD
LVPNGSATAKALDYSLKRWGALTRYLDDGAVPIDNNQVENQIRPWALGRSNWLFAGSLRSGKRAAAIMSLIQSARMNGHD
PYAYLKDVLTRLPTLRSKDISQLLPHQWVQI.

RL060
DNA sequence: (SEQ ID NO: 39)

ATGATCCGCATCGATGCGATCTGGCTAGCCACCGAACCGATGGACATGCGCGCCGGCACCGAGACGGCATTAGCCCGGGT
AATTGCGGTGTTCGGTGCGGCGAAGCCGCACTGCGCTTATCTGTTCGCCAATCGCCGGGCTAACCGAATGAAAGTGCTGG
TGCACGATGGCGTGGGCATCTGGCTTGCCGCGCGTCGACTGAACCAAGGCAAGTTCCACTGGCCCGGCATTCGCCATGGC
TGCGAGGTCGAACTCGACAGCGAACAACTCCAGGCCTTGGTGCTGGGCCTGCCGTGGCAGCGCGTCGGCACAGGCGGTGT
GATCAGCATGCTGTAA

Protein sequence: (SEQ ID NO: 161)
MIRIDAIWLATEPMDMRAGTETALARVIAVFGAAKPHCAYLFANRRANRMKVLVHDGVGIWLAARRLNQGKFHWPGIRHG
CEVELDSEQLQALVLGLPWQRVGTGGVISML.

RL061
DNA sequence: (SEQ ID NO: 40)

ATGCGCCAACGAAGCTCTTACCCGAAACCGTTCAAAGCCCAGGTCGTTCAGGAATGCCTGCAACCTGGGGCAACGGTGTC
CAGTGTCGCCATCAGCCACGGCATCAACGCCAATGTCATCCGCAAATGGCTGACGCTTTATCGAGACCAGCCCGTACCAG
CCTCGTTACCAGCCTTTGTCCCGCTGAAGGCCACCCCTAAACGGCCAGCCGAAACGTCAGTGCTCATTGAACTGCCCATG
GCCGGGCAAATGATCACGGTGAAATAG

Protein sequence: (SEQ ID NO: 162)
MRQRSSYPKPFKAQVVQECLQPGATVSSVAISHGINANVIRKWLTLYRDQPVPASLPAFVPLKATPKRPAETSVLIELPM
AGQMITVK.

FIGURE 30D

RL062
DNA sequence: (SEQ ID NO: 41)

ATGGCTTTATCTCTTATTCGTAGTCTCACTGCGTCCGCCTCACGAAACATCTCGGCGTTGAAACGCGATGCCAAACGCTT
GCAGAAGAACTCCTTTCTTGTGTTTGGAACAGAATATCCACTCAAGGTTTGCCAAAATGCGGTAGCAGTTTCTCGCGGCT
TCCGCTCACTCGCTGATGTCGATAAACTGGAGCAGCACATTGGCATGAATAGAAGCGCTCCATTCTGGGTGATCCGTGGC
CGCAACGATACACACCAGGGGGTACTGGAAGCGCTATATTGTTTAGACCTTGAATATACCGAGAATGGCCCCGTCGTTTT
TACTGGAAACCCAAAGCACTCTATACTTCCAGCCTTAGTCCTTTTTCTTGAGCAAATGAGCTTTAAGAAACTACCCGGAC
TAATCCTCATCGAAACAAAAGAGACCTCAATCCAAACAACCCATATATTCGACGCAATAGAAAAATTAGAAGTCGAAGAA
ACTCTAAATAAATTTCGATTTCTTGACTTGCGAGACCGAAACCTTCCCGTTTCGCTTAGTACCGAGGCTCGTTGCTGGAT
CGAGTCAATTGTCAGTTTATTGCCAAACGACATCCAAGAGGAAATACGTAATAAAGGATGGTCAACTCACTTAGAGATCA
GTGCATATGAGCATGCAAAGTCTCGTAATCAAGTATTTGGCTCCTCCAACTTCCCTTGCGTCCCCTTCCTCTCCATAAAG
TCAGCGATCTATCAACTCATTTCAGGCGCATACCCTCCCTTATGGATGCAGCCATCCTCCTCTGGCGAAATATCTAAAGT
TGATATACGCCGACCTCCTCTCGAAAAAAGCTCAGAGGAAACCTTACTTTATCTCATAAAAAAATTAGAGAATCGACAGT
TCCACACAGGCATTTCATGTGAGCATGAGAGTCGATGGCGGCCGTATGTCGTACTCTTCTCCAGGAATGATCCGGCTAGC
GAGGTACTAGCAGGAGTTATACACTCGTACTTTTCTTGGAAGCAAGATAGAGACCATCGCTCACCCACCCTTTATGTTTC
AGATGGAGCAGTTCCCTATGCTCCCAAGCTTCTAGGTTTAGGCGGCCATACGGTCATTGCAAATGGAATCACTGAAATTC
CCGACGGGGATGGTCTTGGGGAGTTCTATGGCTACAAGAACTCACTTAAAGTCAGCTCCTTATCTAACGGAATACAGTTC
ATGGGTAAGCATGTATCACTAAAGTAA

Protein sequence: (SEQ ID NO: 163)

MALSLIRSLTASASRNISALKRDAKRLQKNSFLVFGTEYPLKVCQNAVAVSRGFRSLADVDKLEQHIGMNRSAPFWVIRG
RNDTHQGVLEALYCLDLEYTENGPVVFTGNPKHSILPALVLFLEQMSFKKLPGLILIETKETSIQTTHIFDAIEKLEVEE
TLNKFRFLDLRDRNLPVSLSTEARCWIESIVSLLPNDIQEEIRNKGWSTHLEISAYEHAKSRNQVFGSSNFPCVPFLSIK
SAIYQLISGAYPPLWMQPSSSGEISKVDIRRPPLEKSSEETLLYLIKKLENRQFHTGISCEHESRWRPYVVLFSRNDPAS
EVLAGVIHSYFSWKQDRDHRSPTLYVSDGAVPYAPKLLGLGGHTVIANGITEIPDGDGLGEFYGYKNSLKVSSLSNGIQF
MGKHVSLK.

RL063
DNA sequence: (SEQ ID NO: 42)

ATGAACGCTCTGACCCAACCGGCCGCCCTCGCCGCCTCCCACCTGAACATCAACCTGACCGACTTCATCGACGAGTTCGG
CGACGAGCTCCTGGAGTCGCTCAATCGCTCCAACCCCCCGGTCTATACCGGCTCCGTCAACGCTCACCGCCAGTTGGTGA
TGGACCGACTCAAGCGCAAGCCCTTCGCGGCCCAGGCCGAGGTCGTCCAGGCCATCACCGCCCTGCTGCTGGACCGTAAC
GAGCAGGCCGGAATCATCAACGCCGAGATGGGCACCGGGAAAACCATGATGGCCATCGCTGTCGCAGCGGTCATGCACGC
CGCCGGCTATCGCCGGACCCTGGTCGTCTCTCCGCCGCACCTGGTCTACAAGTGGCGCCGCGAGATCCTGGAGACCATCC
CAGCCGCCCGCGTCTGGGTACTCAATGGCCCAGATACTCTACTCAAGCTGCTCAAGCTGCGAGATCAGATGGGCGACGCC
TACGACGGGCGCCAGGAGTTCTTCATCCTCGGCCGCGTGCGGATGCGGATGGGTTTCCACTGGCGGCTCGCCTGCTGGAA
GAAACGCGCCGCCGGCGGCCAACTGCTCGCTGCGTGCCCGGATTGCGGACAGGTCCTCGAGGACCTGGAAGGCAACCTGG
TCACGGTGGAGGAGTTCGAGCGTGGTGACCGTCGACGTACCTGTTCCTCCTGCCGTGGGGCGCTCTGGACGCTGATCCGA
CCAGGCAAGCCCGACGGCGGCAACCGGCGCGCAACGATTCTCAAGTCGATGTGCCGGATACCAACCATCGGCCCGGTCAG
GGCGGAGCGCCTGCTGAACGACTTCGGCGAGGACTTCCTGGCCACGATGTTGGTGGACAACGTCTCGGAGTTCATCAACC
TGATGGACGCCAAGGGCAACTTCGTCTTCAGCGATCGGCAGGCCAAACGCATGGAGCGATCGATGGCAAACATCGAGTTC
GGCTTCGGTGAAGGCGGCTACCAACCGACCGAGTTCATCAAGCGCTACCTACCTGATGGCTACTTCGACCTGCTGGTGCT
GGACGAGGGACATGAGTACAAGAACAGCGGCTCGGCCCAGGGCCAGGCCATGGGCGTTCTCGCAGCCAAGGCACGGAAAA
CCGTGCTGCTGACCGGAACGCTCATGGGCGGCTACGCCGACGATCTGTTCTATCTCCTGTTCCGCATCCTCACCCAGCGC
ATGATCGAGGACGGCTATCGGCCCAACGCGCGCGGCAGCATGGCTCCCGCAGCCATGTCGTTCATGCGCGACCACGGTGT
GCTCAAGGATATCTACACCGAGCGCGACGGTGATTCGCACAAGACAGCGCGGGGCAAGAAGCTCTCGGTACGCACGGTGA
AGGCTCCCGGCTTCGGCCCAAAGGGCATCCACCGCTTCGTATTGCCGTTCACCGTGTTCCTGAAGCTCAAGGATATTGGT
GGCAACGTACTGCCCGACTACCAGGAGGAGTTCATCGACGTGCCCATGGCGCCTGAGCAGGCTTCGGCCTATCAGCGCCT
GGCGGCCACGCTGACAGCGGAGCTCCGCCAGGCTCTGGCGCGACGAGATACCACGCTCCTGGGCGTGGTCCTCAACGTGC
TGCTGGCTTGGCCGGACTGCTGTTTCCGACCGGAGATCGTCAAGCATCCGCGAACCCGGGACACACTGGCCTTCGTGCCA
GCGATCTTCGGTGACGAGCAGTTGATACCCAAGGAGCAGGTGCTGGTGGACCTCTGCTTCGAGGAGAAAGCGAAGGGCCG
CAAGGTTCTGGCATACACCGTCTACAGCGGGACGCGCGACACCACGTCCAGGCTGAAGAAAGTGCTCGAGCAATCCGGGC
TGAAGGTGGCAGTGCTACGTGCTTCGGTCGATACCGCTCGACGCGAGGATTGGATCCTCGACCAGGTCGATCGCGGCATC
GATGTGCTGATCACCAACCCGGAGCTGGTGAAGACCGGGCTGGACTTGCTCGACTTCCCGACCATCGCGTTCCTGCAAAC
GGGCTACAACGTGTACACCCTGCAGCAGGCCGCGCGGCGGTCGTGGCGGATCGGGCAGAAGCACCCGGTGAGGGTGGTGT
TCTTCGGCTACGCCGGCAGCTCGCAGATCACCTGCTTACAGCTGATGGCCAAGAAGATCGCTGTGGCTCAGAGCACGTCG
GGAGACGTTCCCGAGTCAGGTCTCGACTCGTTGAACCAGGATGGGGATTCTGTGGAGATGGCGTTGGCACGACAACTCAT
CGCAGCATGA

FIGURE 30E

Protein sequence: (SEQ ID NO: 164)
MNALTQPAALAASHLNINLTDFIDEFGDELLESLNRSNPPVYTGSVNAHRQLVMDRLKRKPFAAQAEVVQAITALLLDRN
EQAGIINAEMGTGKTMMAIAVAAVMHAAGYRRTLVVSPPHLVYKWRREILETIPAARVWVLNGPDTLLKLLKLRDQMGDA
YDGRQEFFILGRVRMRMGFHWRLACWKKRAAGGQLLAACPDCGQVLEDLEGNLVTVEEFERGDRRRTCSSCRGALWTLIR
PGKPDGGNRRATILKSMCRIPTIGPVRAERLLNDFGEDFLATMLVDNVSEFINLMDAKGNFVFSDRQAKRMERSMANIEF
GFGEGGYQPTEFIKRYLPDGYFDLLVLDEGHEYKNSGSAQGQAMGVLAAKARKTVLLTGTLMGGYADDLFYLLFRILTQR
MIEDGYRPNARGSMAPAAMSFMRDHGVLKDIYTERDGDSHKTARGKKLSVRTVKAPGFGPKGIHRFVLPFTVFLKLKDIG
GNVLPDYQEEFIDVPMAPEQASAYQRLAATLTAELRQALARRDTTLLGVVLNVLLAWPDCCFRPEIVKHPRTRDTLAFVP
AIFGDEQLIPKEQVLVDLCFEEKAKGRKVLAYTVYSGTRDTTSRLKKVLEQSGLKVAVLRASVDTARREDWILDQVDRGI
DVLITNPELVKTGLDLLDFPTIAFLQTGYNVYTLQQAARRSWRIGQKHPVRVVFFGYAGSSQITCLQLMAKKIAVAQSTS
GDVPESGLDSLNQDGDSVEMALARQLIAA

RL064
DNA sequence: (SEQ ID NO: 43)

ATGGCCCTCATGTTCCCGCGCTTGGCGCGCAACTTTGCACGCAACGGCTACTTCCCTACCGATGAGGTC
ACCCTCGAACGCGCTCTGCAGGCCCTCACTCTTGCCCCGTCGGGAAGGATGAGGATCTGTGACCCCTGCGCCGGTGAGGG
TGTTGCCCTGGCTGAGGCAGCACACACCCTCGGCCGCGATCAGGTCCAAGCCCTCGCTGTCGAGTACGACCGCGAGCGCG
CCGACCATGCCCGAGGATTGCTTGACCGAGTGCTGCACAGTGACCTTTTCGACACCATGATCAGCAGGCAGTCGTTCGGA
CTGCTCTGGCTCAACCCGCCTTATGGCGACCTGGTGGCGGACCACTCCGGTGCGTCGCAGTACCAGGGCAGCGGCCGCCG
GCGTCTGGAGAAAGCGTTCTACCAGCGCTGCCTGCCGTTGCTGCAGTACGGCGGCGTCATGGTTCTGATTGTTCCTCACT
ACGTCTTGGACGATGAGCTGACTGGCTGGTTGAGCAACCACTTCACCGGCCTGCGCATCTACGCAGCCGCGGATCCTACC
TTCAAGCAGGTGGTGATCTTCGGCATCCGGGTCCGTCGGCAGGACCTGGCCCGGGCGGACGCCAATCAGGTGAGGTCTCG
CCTGCAGGCGATCGGAGCGGGCCAGGAAAAGGCCGAGGAAATTCCAGCGGCTTGGCCGTGGGAACCCTACGTGGTTCTGC
CGGCCACCAGCGAGCTGGAGCACTTCTACCGAGTAACCCTGGAGCCGGAGCAGTTCGCCGGTGAAATCCAGCGGCTGCGA
GGTCTCTGGCCTGACTTCAACCTGCACTTCGCGCAAGCGGGGCTGCAGCCGCGCCCACCAGTCCGCGAGCTGTCTCGCTG
GCACCTGGCCCTGGCCCTGGCCGCCGGCGCGATATCTGGCGTCGTGCGATCGAAGTCGGGCCGGATCCTGGTCGTGAAGG
GTGACACCTACAAGGACAAGGTCCGCAAGACCGAATTCACCGAGGACGACGACGGCAACATCACCGAGGTGAGGATCCTC
ACCGACCGTTTCATCCCGATCATCCGGGCATGGGAAATGACACCCTCCTCGGTCAATCAGGGCCGCGTGCTGACCATCAG
CTCCTCGGCCGCGACCACGGAAGAAGCTGAAGAGCCCCAACCTGAGCCGGCCCCCGCACCCGCACCGCTGCTGATCAGCC
CTGGCCGGGTCGTAATGACCGCAGCCGTGAGCCACCTGGTGGAAACCGGTCAACTCAACCCAGCGCCTTTGCTGAAACGC
CATCTGGCGGGAGATTGGGGAACGCTGGACCAGGAAGACTGGAACACCAACCAGAGAGCCCTGAAGTTCGGCGATCGGCT
GCTGTCGTCCTACGACATCGACGCCGGCGACGAATCCAGGCTCTGGATCATCACTGAGGCAGACCGCAGCTCAACCACGC
TTTTGCTCCCTAGCGATTACTGA

Protein sequence: (SEQ ID NO: 165)
MALMFPRLARNFARNGYFPTDEVTLERALQALTLAPSGRMRICDPCAGEGVALAEAAHTLGRDQVQALAVEYDRERADHA
RGLLDRVLHSDLFDTMISRQSFGLLWLNPPYGDLVADHSGASQYQGSGRRRLEKAFYQRCLPLLQYGGVMVLIVPHYVLD
DELTGWLSNHFTGLRIYAAADPTFKQVVIFGIRVRRQDLARADANQVRSRLQAIGAGQEKAEEIPAAWPWEPYVVLPATS
ELEHFYRVTLEPEQFAGEIQRLRGLWPDFNLHFAQAGLQPRPPVRELSRWHLALALAAGAISGVVRSKSGRILVVKGDTY
KDKVRKTEFTEDDDGNITEVRILTDRFIPIIRAWEMTPSSVNQGRVLTISSSAATTEEAEEPQPEPAPAPALLISPGRV
VMTAAVSHLVETGQLNPAPLLKRHLAGDWGTLDQEDWNTNQRALKFGDRLLSSYDIDAGDESRLWIITEADRSSTTLLLP
SDY.

RL065
DNA sequence: (SEQ ID NO: 44)

ATGCCCAGTCCCACCCCGCTCTACCAGATCGAAGAGTGTCCAGACCTGTACGTCGACGCCTGCGTGTGCGACGAGCAGTG
CAACCTGGTCTTTCTTTCGGCCTGGGGCCGCGACACCGTGACACAAGAGTTCCTGGCCAGGCTGACGCTGGGCCGGGAAG
AAAATGGCATCGACCATTTCCACATCATCGTGGACGGCCGCCGCTTACCTGTCTTCCCAAACCAGGATCTCCTGGAGAAA
CGCACCACCCGTCAGTTCCGCGGCACGTTGTTCGGCAGCCTGCTCAATCTTTGGCTGTTCGATCGGCGCGCCTCGGCGCC
CGACCGAGGCAATCACCTCGCCTTCGCACTCCTGCAGCGCGATGAGGATCCACACCAGAGGCTCTGGCCGCTGGTGATGG
AAACCTGTCCGCTCCCCCTCCTGCAGCACTGGCGCGAGCCGGTGATGGAGGTTCTCACCCAGCACCAGATGTTGACGGCC
CTACCCGGGACGATCGGCAACGTCTGCGCCTGGCGACTCGCCCTGCGGGTCGACGTGCTTGAGCCCACCCTCGGTGAGGT
AATCCGCGAAAGCATTCTTACCACCGATGCTCAGGCGCAAGCCTGA

Protein sequence: (SEQ ID NO: 166)
MPSPTPLYQIEECPDLYVDACVCDEQCNLVFLSAWGRDTVTQEFLARLTLGREENGIDHFHIIVDGRRLPVFPNQDLLEK
RTTRQFRGTLFGSLLNLWLFDRRASAPDRGNHLAFALLQRDEDPHQRLWPLVMETCPLPLLQHWREPVMEVLTQHQMLTA
LPGTIGNVCAWRLALRVDVLEPTLGEVIRESILTTDAQAQA

FIGURE 30F

RL066
DNA sequence: (SEQ ID NO: 45)

ATGAATCCATTGTTCACCAACCTCACCCAGGAAACCCTCGCCTACCTCGAGGACCAACTGTCCAACAACGACGTCGCCGG
CGACGACGAGCTCATCGACTTGTTCATCGAGGAGCTGTCGCTGACCTTGGAGCAGGCGGAAGCGGCTGTCGCGCTACGCG
ATCAGTACCTCTGCCAGGTCTTCCTGATCGGCCAAGGGCCGCTGCACCAAGCCGATGGACTCAGCTTCGACCCTCACACC
AAGAGCGTTCGGTAG

Protein sequence: (SEQ ID NO: 167)
MNPLFTNLTQETLAYLEDQLSNNDVAGDDELIDLFIEELSLTLEQAEAAVALRDQYLCQVFLIGQGPLHQADGLSFDPHT
KSVR

RL067
DNA sequence: (SEQ ID NO: 46)
ATGGGATGGCTTTTCTCACATCAGACGAAGGAAGACCTGCTGCGTGAGCTGCTGGCCCCAACCAGTACCTTCGCAGGCAG
CACCGAGGTGCTGGCACACGCAGTCTCCGGCAATGAACTTTGGACTGTCGTAAAACGAACTTTTCACCTTGCCGGATTCT
ATTTCGGCAAGCCGGCCGGTCACTCGATCACCATGATCGAGCTGCACTTGCTGGACTGCTCGGCCGGGCAATGGGGCTAC
AAGACCATTCCGGAAAGCGCCGGCCCGTTCTACTACGGCTGTCCGCTGGAGTTCCTGGACCTGGCTCACGATGAGATCAA
CCAGGAATGGCGTAAACGCCTGACGCACGAACACCAAGCCTGA

Protein sequence: (SEQ ID NO: 168)
MGWLFSHQTKEDLLRELLAPTSTFAGSTEVLAHAVSGNELWTVVKRTFHLAGFYFGKPAGHSITMIELHLLDCSAGQWGY
KTIPESAGPFYYGCPLEFLDLAHDEINQEWRKRLTHEHQA

RL068
DNA sequence: (SEQ ID NO: 47)
ATGAAATCGATCTACAACACCCCAGGCTTCAGCGAGGAGTTGTTGCTGGTTTGCGCCTCGCTGCGCGAGGTCGGACTGGA
CAATCTGGCTGACCAGTTCCGCGCGGCAGTGTTCGACCGATCCGTCGTCGACCAGGCCATCATCGCACTGCGTGAGCGGG
TGAAGACCCCTTCGCCGGAGCATGCGGCCGACAACGAGCCCTGGTTGTACTGCGACTGGCAGGCCAGGCAAACAGCTTAC
CGGCTCCTCCAGCGCCTTGAGCGCGCAACACGCTGA

Protein sequence: (SEQ ID NO: 169)
MKSIYNTPGFSEELLLVCASLREVGLDNLADQFRAAVFDRSVVDQAIIALRERVKTPSPEHAADNEPWLYCDWQARQTAY
RLLQRLERATR

RL069
DNA sequence: (SEQ ID NO: 48)

ATCCCCTCACCACGATCCCGCTTCGGCGGGATCATCCTTTTCGCAGGTCATACCATGATCACAGTTCCCGGACAGTTGGC
CATTCGAACCATCAACGGTCGCTATGGCGAGTTCAATGTGGGAAAACTCTGGACTTCGATCGGGGAGTTCATCATCAAGG
ATGCCTTCCTGGATCAACACACCGAAGGCAAGTACCGCGGTGATTTCGTCATCGCCAATATCCGCCCCCACCACTACTCC
GCCGGCGGTCGGCTAGTCATCGAGATCCGCGCCATAGTGGACAGCATGACGCTGAACGATATGGACAGCCTCAGCGACGA
GGAGGTAGAGCGTCTTTCCGGCAATGAGGTGGATCCGCTCGACGAAGTGCCCGAGATCCAGCTCCCCACAGTAGTACCGG
CGATACCACCAAAGTCGCCGTCACCCCAGAAGTCGAAGCCTCTGTGCCTCGCTGCAACCAGGGACGCGCCTTTCGGTATG
GACACTCCGGCTCCTGCAGAGCAGGCCGCCTCTCTGGACACAGACGCGGATGCAGAACTGTTCGGGACGGTCTGGCCGCT
AGGCGAAATCGTCAAGCTGGACACCACGGTCGACCGCAAGCGACTACGCCAACAGTGCGTGCGACTCGGCGCGCTGGGCT
ATGAGCTCGACTTCAAACAACAGGTGTGGACCCGCAAGGAGGCCGCATGA

Protein sequence: (SEQ ID NO: 170)
IPSPRSRFGGIILFAGHTMITVPGQLAIRTINGRYGEFNVGKLWTSIGEFIIKDAFLDQHTEGKYRGDFVIANIRPHHYS
AGGRLVIEIRAIVDSMTLNDMDSLSDEEVERLSGNEVDPLDEVPEIQLPTVVPAIP

RL070
DNA sequence: (SEQ ID NO: 49)
ATGACCTCTCTCAACAACCACTCCAGCGCAGGTCACACTGCTGCGTACCTCAAACTCCCGATCGTTCTCACCAACGCGGC
CTGGCTGCGCCTGGTCTATCTCGCCAACCCTGCCAGGGTCGACGAGATGGGCACCCGGCTGGCCAGTGTCGTTCAAACCG
CCTGGCAGGAGCTTTCTCTCCAGCCGACCGCGAAGCACATCCAATTCCACCTGTACCACAAGGAGGAAGAGGGGCAGGAC
CGCGCGCTCGCGCTGCTGGTTCTCTCGATAGTCGAGCCGTCCGATGAGCCTTCCTACCTGCGCATCGAGTTGCAGGAAGA
GTGCCTCGCCGAACACCCGGTTACCGAGTAG

Protein sequence: (SEQ ID NO:171)
MTSLNNHSSAGHTAAYLKLPIVLTNAAWLRLVYLANPARVDEMGTRLASVVQTAWQELSLQPTAKHIQFHLYHKEEEGQD
RALALLVLSIVEPSDEPSYLRIELQEECLAEHPVTE.
PKSPSPQKSKPLCLAATRDAPFGM
DTPAPAEQAASLDTDADAELFGTVWPLGEIVKLDTTVDRKRLRQQCVRLGALGYELDFKQQVWTRKEAA

FIGURE 30G

RL071
DNA sequence: (SEQ ID NO: 50)
ATGACTCAACTCAACCCGTTTATTCGCGGCTATGAGAGTTTCCGCATCGAGCGAAACCTGCAGATCACTGACGAAGGCAA
CAATCTACCGTGCTACCGCGCTCTGCATGAAACCCAGCAGCACCTCCCAGACGAATATTTTCAGTGCGAGCTGTGCTACT
TCAATAACGATTTCGCCGTGGTAGTCCAAGAGTTAGACGATGAAAGAGTTGAAAAATGCCCTCACCAAGGAATAGTGAGA
AACGTACTTTACAGCATCTACGGTGAGCAGGACGGCAGAAAAAAGCTTATCGGAGATCAATACTCACTGACCGAAGCCGA
GAGTGTCGTTCGATACCTTTCGTTCGGCGGCGGTTATAACCCCTGCTGGGAGATCAGAAAAACACATCTACCCATCAGCG
CGTGGAATAGCCTCTACGAAAGGTTCTCGACCAAGATGCCAATCCGCTTGCCCTCGGTGTTGGTATCGCTCTTCTGGTGT
AACGAGCACGGTGCCGTGGGCTTTCGCTTGCACAACACCCCTTGGACGGATGAGTGTCTGGAGATCCTGGAGATGACCGC
AGCCGCTCTTCGACAAGAACAGCTTGCCTTCGGCCTCGACGAACACCTTGTCGATCTGCTTCACCTCGCGGGACAAGCAG
ACATTCGGCTCCTGGTACTTGATCCATTCGCGCCCACGCTCAAGGGCCTGCCGCTTTATGACGATTGA

Protein sequence (SEQ ID NO: 172)
MTQLNPFIRGYESFRIERNLQITDEGNNLPCYRALHETQQHLPDEYFQCELCYFNNDFAVVVQELDDERVEKCPHQGIVR
NVLYSIYGEQDGRKKLIGDQYSLTEAESVVRYLSFGGGYNPCWEIRKTHLPISAWNSLYERFSTKMPIRLPSVLVSLFWC
NEHGAVGFRLHNTPWTDECLEILEMTAAALRQEQLAFGLDEHLVDLLHLAGQADIRLLVLDPFAPTLKGLPLYDD.

RL072
DNA sequence: (SEQ ID NO: 51)
ATGGGACTGGTGTTTCCTACCGAAAGGAGAATCACCATGCAATACGGAAAGCTGGCGCTCGCCCATCTCAGCCTGGAACT
GCCGTTGCAGGTACTTATGAATAAGAACCGTGCTTACTACATCGGCACTTCTGACGAAGAAGGACCAGCCTCGCGCGAGT
CGGTTGAATATTACCCCTCACGCGAACTTGCCCAACAGGCATTAGACCACGGCACTTGGACGCAACTGGAATATTAA

Protein sequence: (SEQ ID NO: 173)
MGLVFPTERRITMQYGKLALAHLSLELPLQVLMNKNRAYYIGTSDEEGPASRESVEYYPSRELAQQALDHGTWTQLEY.

RL073
DNA sequence: (SEQ ID NO: 52)
ATGGGAAATGTTTGGCGATTATGCCAGGGCAGATACCTGGGCATTGTTGTTGGCCAGGAACAGCCAGGCGAAGTTGCAGA
ACTGACTGCTGAGCAGCAGCTCGTCCTCGACGTCGCTGAGGCTAACCTCCTCAACTTCCGGCAGGGCGGGCAGTTCTACG
ATTTGGATGTTGCTCATGATGATCTCCAGATAATGGAGAACACCACGCCCTGGGGGGAGATGGTGCCCCCCGGATGGGTA
TGCGATGAAGAGTGGCGCATAGCGTAG

Protein sequence: (SEQ ID NO: 174)
MGNVWRLCQGRYLGIVVGQEQPGEVAELTAEQQLVLDVAEANLLNFRQGGQFYDLDVAHDDLQIMENTTPWGEMVPPGWV
CDEEWRIA.

RL074
DNA sequence: (SEQ ID NO: 53)
CTGACGGGCAAGGTGTTTCTCCGCTTTCGCTTACGAAACTGGAGAATCATCATGAGCAACAACACCCAAGCCCAAGAAGC
CAAGTATTTCGACCTGCACACCACCGGTATCGGCTACCTCAATCGCATCCGCGAGGTACCGATCCGCCGAGGTGAACCAT
TCCTCGCCGTAACCGTCGCAGCCCTCCATGGCGCGGCAGACAGCGTGGAATACTCCTACATCGACTGCAAAGTGGTCGGC
GCCCAGGCTGAAAAGCTTGTCCGCCGTTGCAAGGAAGCAGTCGAGGCCAAGAAGAAGGTTCTGATTTCCTTCCGTATCGG
CGATATCTGGGCGGATCCCTTCATCCACCAGAAAGGCGAGAAACAAGGCAAGCCCGACGCAAGCCTCAAAGGCCGGCTGC
TCTTCATCTCCTGGATCAAAGTGGATGGCACCACCGTCTACGATGCGAAGGAAGAAGCTGAAAAAGCCCAGCAAGGCAAA
GGCGAACCTCAAGGTGAGCCCGCAGCCCCCGCTGAGCACGCTGAACAAGCCGCTGCTTGA

Protein sequence: (SEQ ID NO: 175)
LTGKVFLRFRLRNWRIIMSNNTQAQEAKYFDLHTTGIGYLNRIREVPIRRGEPFLAVTVAALHGAADSVEYSYIDCKVVG
AQAEKLVRRCKEAVEAKKKVLISFRIGDIWADPFIHQKGEKQGKPDASLKGRLLFISWIKVDGTTVYDAKEEAEKAQQGK
GEPQGEPAAPAEHAEQAAA.

RL075
DNA sequence: (SEQ ID NO: 54)
ATGTCCAAGCAATCCACCAGCTTCGAAATCGGCTTTGCCCTCGGCAGTGTTGTGCGTGAGTTCCGCAGAGCGCTCAGTCG
CCCTCCGGTCGTAGTGCAAGCACAAGCGCCGGTTGCGTTGAGAGTCCAGCGCATCGATCCTGCCTTCCTGGCCGGCCCGA
CCGCTGGCGAGCTAGAACACATCAGCGACATCCCAGCCATCGTCCGGCTGAAGAAGGTCAACCTGAATGACTGGTATCTA
GCCAATACGCGCGAGGTGCAAAAGCCCAAGCGCGCACGCAAACCCAAGCCGGCCAAGGCGACCGCCAAAGCTGAAACGCC
AGTCAGGAAGGAGCTCAAGATGGGTTCCCTCGACCATTTGATTGCACCCAACTCCGAAAGCGAAATGGGGAGGCCCCCTC
TCCAGTTAGAGTCCCTGAACGATCATGAGATTGCTCTTTTGCCAGCACCTCCTGGTAGCGCAGTCTCTTGGGAACTCCAT
CGGCGTACTCAGGAGCAATACCAACAACGCTGGCAGGACTACTTGTCCACCATGACGGATGAACAAGTAGCTGCTCTCGG
CCGCTAA

FIGURE 30H

Protein sequence: (SEQ ID NO: 176)
MSKQSTSFEIGFALGSVVREFRRALSRPPVVVQAQAPVALRVQRIDPAFLAGPTAGELEHISDIPAIVRLKKVNLNDWYL
ANTREVQKPKRARKPKPAKATAKAETPVRKELKMGSLDHLIAPNSESEMGRPPLQLESLNDHEIALLPAPPGSAVSWELH
RRTQEQYQQRWQDYLSTMTDEQVAALGR.

RL076
DNA sequence: (SEQ ID NO: 55)
ATGGTGTTTCTCCTGCAGGTTGAGGGTGCGGAGAAAACACTGGCCCTGGCGGGGAAGTGGATTCCCCGCTGGGTTGCGGA
AGGGAGCTTCTATCGACCGAGGCCGACCGACCGCGCTACCAGAAGCTATGCGGTCTTGGGTTGGATCAACACGGTGGGCT
GTGCTGCAGCATTTCGGATCCGAGCTGCATGGGGGCATGTCGCTGACAACGTCAGCAGATCACGCGTTCATCATCGAAGC
GGGGGGCGAAAGTGTCAAGGTCAAGCAGGAGGGGGAGCGGATGCAGCGGGCGGAGAGCGAGGGCGGAAGAGCGCGGCTGG
TAGAAACCCTGTCAAAGGTTTCCCCAGCCGTGTCTGGAAGGGGAGTCAAGTGAGCCACCTGTGGTTGAATCGTCGATCCC
TGGGCATTGATCGTCTCGATCCCATCACCCGGCCATTATCGTGGCTTGGCCAGCAAACAGTAGGCACGCATCCGCGTACA
AAGGGAGCCCTGCGTATCACCGGCGGGCCACCGGCAGGGAGAAGGATCCCGATGGGTAGCCTGATAGTCCTGGAGCAGGA
GCATCAGGCTACCCATGGAGAGGGGAAAAGGAGGGGCCGTAACACCAGTACGACCCTTAAATCGAGGAAACACCGAACCT
CTTGA

Protein sequence: (SEQ ID NO: 177)
MVFLLQVEGAEKTLALAGKWIPRWVAEGSFYRPRPTDRATRSYAVLGWINTVGCAAAFRIRAAWGHVADNVSRSRVHHRS
GGRKCQGQAGGGADAAGGERGRKSAAGRNPVKGFPSRVWKGSQVSHLWLNRRSLGIDRLDPITRPLSWLGQQTVGTHPRT
KGALRITGGPPAGRRIPMGSLIVLEQEHQATHGEGKRRGRNTSTTLKSRKHRTS.

RL077
DNA sequence: (SEQ ID NO: 56)
ATGCCGCTGATGTGGATCGTCCTGGTGCTCGCGCTCATCACCGGGACCTGGCTGAGTGTACAAAGCGACCACGCGACCTC
GAGCGCCGAACTGGCCGAGGTCGACACCCTGGCCAGGAGCTTGCTGCTCTTCCGGTCCAGTCTGGCGGAGTACGCACACG
CCAACCCCGGTTTCACCGGTTCGCCGGCGGACTCCGCTCTTGGTTTACCGGCCTGGTTCCGCAAGCCAGCGCGGCTTCAG
GGCTACATCGCCGCCGGCACCAGCTACGCCTTCATCGCCTCGCCGCCGGCGGGGCTGGCGGCGGCCGTGGATGCTGGTAC
GGAATCCGACCTGGTTGGCGTCAGGCGCAACGGCCAGTTAGTCACGCGCCGCCTCGGAGCCACTGTCATTGCGCTCCCTA
CGCCCATCCCCGAGGGCGCGGTGGTCGCGGTCAAATAA

Protein sequence: (SEQ ID NO: 178)
MPLMWIVLVLALITGTWLSVQSDHATSSAELAEVDTLARSLLLFRSSLAEYAHANPGFTGSPADSALGLPAWFRKPARLQ
GYIAAGTSYAFIASPPAGLAAAVDAGTESDLVGVRRNGQLVTRRLGATVIALPTPIPEGAVVAVK.

RL078
DNA sequence: (SEQ ID NO: 57)
ATGAGGAGTACGCGCAGCAGTGGATTCATCTCGATCGAACTGATGATCGCCCTCGTCGTGATCGCCATCGCGACCGCCGG
TGGCATATCGGTCCTGATGAGCTACCTGGACGGCTTGGACGAGCAGCACGCGGCCCAGCAGCAACAGCAGGTGGCCAAGG
CAGCGGAGAAGTACCTGAAGGACAACTTCAGCACGGTTCTGGCCAGCGCCGGCGCCACGGCCCCGGCGGTGATCACCGTC
CCGATGCTGCGCAACACCCGTTACCTGCCCGCAGGCTTCCGCGACACCAACATCTACGGCCAGCAATACCAGGTCCTGGC
CCGCAAGCCGGCGGCCAACCAGCTCGAAACGCTGATCGTGACCACGGGTGGACAGGTAGCTTCCGAACTCTCGATCCGCC
GGATCGCGCAGCTCATGGGAGCCACCGGGGGCTACATCTCGAAAACCAACACCAGTATCGCCCAGGGCGCCGCCTGGCAG
GTGGCCTTAAGCAATTTCGGTAGCGCTCCCGGCGCTGGACATCTGGCGACGGCGCTGTTCTTCCAGGACGGCGCCATCGC
CAACGAGTACCTCTACCGCAATGCCGTCCCGGGTCATCCTGAACTCAACCGGATGAATACCACGCTGGACATGGGAGGCA
ACAATATCGCCGCAGCCGGGGCGATCACGGCCAGCGGCAACATCACCACCAGCGCGGACATCAGCGCGCGCAACGTGACA
GCCACTGGTACGGTGAAAGCCGGCACTGCTGACGTCGCCGGCGAGACGTACACCGGAGGCTGGTTCAGGACCCGTGGTGA
CACGGGCTGGTACAACGAGAAATGGGGCGGCGGCTGGTACATGAGCGACAGCACCTGGGTGCGCTCCTGGATGAACAAGA
ACGTCTACACCGGCGGCGAGATGAAAGCGGGCAAACTCACCGCCGAGGGCCGGACGGAAGTCGGCGAGTACCTACAGCTC
AAAGGCGTGGCCACCGAAGGAGCCAACTGCTCGCCGAACGGGCTGGCAGGCATCACCAGCACCGGACTCTGGCTGTCCTG
CCAAAACGGGAAATGGGGACGAACCGCCGCCTCCATGCGCCTGAACACCACGGCCGGCGTGATCAAGGACTGGTGTACGT
TGCATGGTCAGGATAGCGCCATGGTGAACTACGACTACGTCCGCTACGCGATCACCTGCGGCGGCCGATTCTGCGCAGTG
GGCTTCAACCAGACATTTGGCACCAACTACTCGTTCGGGCTAATCACTGAGATCGGCCCAGGCTTCAACTACCCGGAACC
CTACAAGACCCCCGACTCGACCAACGTGACCGTTACCTGCGTGAACTAG

Protein sequence: (SEQ ID NO: 179)
MRSTRSSGFISIELMIALVVIAIATAGGISVLMSYLDGLDEQHAAQQQQQVAKAAEKYLKDNFSTVLASAGATAPAVITV
PMLRNTRYLPAGFRDTNIYGQQYQVLARKPAANQLETLIVTTGGQVASELSIRRIAQLMGATGGYISKTNTSIAQGAAWQ
VALSNFGSAPGAGHLATALFFQDGAIANEYLYRNAVPGHPELNRMNTTLDMGGNNIAAAGAITASGNITTSADISARNVT
ATGTVKAGTADVAGETYTGGWFRTRGDTGWYNEKWGGGWYMSDSTWVRSWMNKNVYTGGEMKAGKLTAEGRTEVGEYLQL
KGVATEGANCSPNGLAGITSTGLWLSCQNGKWGRTAASMRLNTTAGVIKDWCTLHGQDSAMVNYDYVRYAITCGGRFCAV
GFNQTFGTNYSFGLITEIGPGFNYPEPYKTPDSTNVTVTCVN.

FIGURE 30I

RL079
DNA sequence: (SEQ ID NO: 58)

GTGAGTGTGAACCCGATCATCCAGGCTCAGTTCGTCGACCTCTACCTCGGTGAAGGCTTCGCCGACGTGAAAGGCCTGGC
CGGCGCCGGCGCGCGCCGAGTCGAAGTGCCTCGCGAGTGGGAGTCGCACGTCCAGGAACTGCTCCAGATCTGCAGGCAAA
CGCTGGAGGAGCTGCAGGACCCTGAGTTCGCCATCGTCGTCGACGGCGTTCTGCTTCGCGTCACCCTCCTCGAAGACGCT
TTCAGTGGCAGCGTCTTCGTGCTGCGCCGGTCGAGCGCCCAATTGCGGGAGTTCCAAGAGATCGGCTATCCGAGCGAAGT
GGTTTCCGCACTGATGGATCCGCAGTTGCAGGGCCTGGTCCTGTTCTGCGGCGAGATGGCGACAGGCAAGACCAGCTCCG
CCGCCTCTCTGCTCCTGGCCCGCCTGCAGGAGTTGGGCGGGGTGGGCTGCGCCGTCGAGGACCCGCAGGAAACCAACCTC
AGCGGTCAGCATGGGCTCGGCCGCTGCATCCAGGTCAGAACCTCACGGCGCTCAGGCGGATACAGCGAGGCCCTGCTGCG
CACGCTGCGGGCCGGCGCCGACCTGGTGCTGATTGGCGAGATCCGCGACGAGGACACCGCCTACCAGGCCTGCAAGGCCT
CTCTGACCGGCAGCCTGGTGATCGCCACCATTCACGCGAAAAGCTGTCATCAGGCGATCGAGCGCTTGGTGACGCTCGCC
CAGCCACTGGCGAGAAACGCCTACGACGTGGTTGCCGAAGGCATCCAAGCTGTGATCTGCCAAGCGCTGGAGAGCGATGG
TTCCTCGCGCCGCCTGACCGCCGAGCCACTGCTGTTCACTGGCGACGACGGCCCGTCCATGCGCGACAAGATCCGCCGAA
AGGAGGCTCATCTGCTGCAGGACGACCAAGCTCGCCAGTCCCGGCAAAGCCTATGGAGATAA

Protein sequence: (SEQ ID NO: 180)
VSVNPIIQAQFVDLYLGEGFADVKGLAGAGARRVEVPREWESHVQELLQICRQTLEELQDPEFAIVVDGVLLRVTLLEDA
FSGSVFVLRRSSAQLREFQEIGYPSEVVSALMDPQLQGLVLFCGEMATGKTSSAASLLLARLQELGGVGCAVEDPQETNL
SGQHGLGRCIQVRTSRRSGGYSEALLRTLRAGADLVLIGEIRDEDTAYQACKASLTGSLVIATIHAKSCHQAIERLVTLA
QPLARNAYDVVAEGIQAVICQALESDGSSRRLTAEPLLFTGDDGPSMRDKIRRKEAHLLQDDQARQSRQSLWR.

RL080

DNA sequence: (SEQ ID NO: 59)
ATGAGCACTACGCAACGCACTTCCCGTCCGACGCAGGGCGGTTTCGTTTCCATCGAGATGATCATCGTGCTGATCATCAT
CGCCATCGGGGTCGGCCTGGGCCTGGCCGCAGCGGCTGGAATGTTCAGTTCGTCCAACGCCAACGAGGAACAACGCAACA
TCAGCGTCATTGCGGCCAACGCACGCGCCCTGAAGACCTCTTCGGGCTACGGCTCCAGCGGTACCAACCTGATCCCCAGC
CTGATCGCAATCAACGGCGTGCCGAAGAACATGAGTGTCTCCTCCGGCGTCGTCTACAACGTCTACGGCGGATCGGTCAC
TGTCTCGTCCACCGGCATGGGCTTCTCGATCACCACCAGCAAGTTGCCCCAGGACGCCTGTATCACGCTGGCCACCAAGA
TCGCGAAGAACACCTTCGAGCAGACCAAAATCAACAGCGGATCCTCGATCACCGGAGAAGTGACCACCGCAGCCGCGACC
CAGGCCTGCAGCAGCGACAGCAACAGCATTACCTGGACCTATAGTTCGTGA

Protein sequence: (SEQ ID NO: 181)
MSTTQRTSRPTQGGFVSIEMIIVLIIIAIGVGLGLAAAAGMFSSSNANEEQRNISVIAANARALKTSSGYGSSGTNLIPS
LIAINGVPKNMSVSSGVVYNVYGGSVTVSSTGMGFSITTSKLPQDACITLATKIAKNTFEQTKINSGSSITGEVTTAAAT
QACSSDSNSITWTYSS

RL081

DNA sequence: (SEQ ID NO: 60)
ATGGGGGGCTTCTGGGAGCAGTTGCAGTTCGCCTTCTACAGCAAGCAGTTCGGCCGCAAGGAACGCCTGCAGTTCTACGA
AAGCATGTCCACCCTGCTCGAAAACGGGGTCCCGTTGAAGGATGCTGTGGCAGAGGTGCACAAGATCTTCGCTCATGAGG
GGCAGCATCCGTTTCATCCGGTGGCCATCGCCAGTCGCGAAGCGCTGATGGGGCTGTCCAACGGCAAGCGTCTGGCCACC
GCCATGGCGCTCTATCTCCCCGCCCAGGAGCGAGCGTTGATCGAGGCCGGCGAGATGAGCGGCAACCTGGTTCAGGCCAT
GGGCGATGCCGTCTCCCTGGTCGAGGCCCAGGCCAGGATCCGCGCCACCATCTGGCAGGCGCTGCTCTACCCCTCGGCGC
TGTCCGCCATGATGGTGTTCCTGCTGTGCATCGTGGCCTATCGCATGGTCCCCAGCCTGGCCAGGCTCTCCGACCCAGTC
ACCTGGACCGGCCCGCTCGCCACGCTCAACGCCATTGCCAGCTTCGTCACAGGACCTGGTATCTACGTTCTGGTCGCCGT
CATCACCCTCACGGTGGTGGTCATCGTCACGTTGCCGACCTACCGCTGGAAAGGCCGGGTCTGGCTGGACCGGACGCTGC
CGCCCTGGTCCATCTACCGCATGCTCCAGGGCACCACCTTCCTGCTGAACATGGCGGTCATGCTCAACGCCGGCATACGC
CCCTACGACAGCCTGGCCAGCATGATCAAGATCTCCCCGCCCTGGCTGAAGCAGCGCTTGGAAGCTGCCCGCTACGGCGT
GGGCCTGGGCCAGAACTTGGGTGTTGCCCTTCGCAGCGCCGGTCACGATTTCCCCGACCGGCAGGCCATCCAGTACCTGT
GCATCCTCGCCAACCGGGGAGGCTTCTCCGAGGCGCTGGTCAAGTTCAGCCGCCGCTGGCAGGAGACCAGCCTCAAGCAG
ATCGAGCTGGCCGCCGGGCTGGTGAAGAACTTCGCCCTGATCTTCATCGGCGCGCTGATGATCCTGGTCCTGCTCGGCGC
CTACCAGGCACAGCAGCTCATCCAATCCATGAACCACTGA

Protein sequence: (SEQ ID NO: 182)
MGGFWEQLQFAFYSKQFGRKERLQFYESMSTLLENGVPLKDAVAEVHKIFAHEGQHPFHPVAIASREALMGLSNGKRLAT
AMALYLPAQERALIEAGEMSGNLVQAMGDAVSLVEAQARIRATIWQALLYPSALSAMMVFLLCIVAYRMVPSLARLSDPV
TWTGPLATLNAIASFVTGPGIYVLVAVITLTVVVIVTLPTYRWKGRVWLDRTLPPWSIYRMLQGTTFLLNMAVMLNAGIR
PYDSLASMIKISPPWLKQRLEAARYGVGLGQNLGVALRSAGHDFPDRQAIQYLCILANRGGFSEALVKFSRRWQETSLKQ
IELAAGLVKNFALIFIGALMILVLLGAYQAQQLIQSMNH.

FIGURE 30J

RL082
DNA sequence: (SEQ ID NO: 61)

ATGACGAACCTTCAGATTGCCGCGCTTGCGCAGCCCTCCATGGTGACCCAACTGCTCACCGCCGACGGTGGTGAATGGGA
GGTATCGAAGCACCTGCAGGAAATCATGGCTCTGGCTGCCGACGGCACGCTCTATCTATCGGAGAGCCACCAGAACGACA
TACACGTTCTGTCGTTCATCGACCGTCTCGATCGCCGTGGCTTCCGATACCAGCTCAACCTCACCGACCTGCAGACCATT
CACCAGCTTTACCGCGCCGTCGCCATGGACGGCCTGGTCGATAGCGATGGCCAGCGCGCCACCCAGATGCAGGAGCGCGT
GGTCAAGATCATTCGTAAGGCCACTGAGCTGCGCGCCAGTGACGTGCATTTCGTCGTGAGTCCCGCCGGCACCGGCAGCA
AGATCCGTTTCCGCGTCGACGGCCTGCTGAAGACCGTCGAGCAGTTCCGCAGCCAGGAGCTGCACGAACTCTGTGCAACC
ATCTACCAATCCATGTGCGACGTGGCCGAGCCACTGTTCAAGCCGCAACTGGACCAGGACGCGCGGATGAGCCAGACCTT
CGTCGAGAAGCTCAACCTGTTCAGTGCCCGGATCGCCACCCGCCCGCGTGCCGGGGGGTTCCTGATGATCCTGCGACTGC
TCTACGACGACACCGGCCTCGACAGCCTGGAGCAGCTCGGCTACCTGCCCGAGCAGAACGCACTGTTCGATCGCATGATG
CGTATGCCCTACGGCATCAACATCCTGTCCGGCCCCACCGGGTCAGGAAAGTCGATGACCTTGAAGGTCACCCTGGAAGG
CCTCGACAAGCTCCATGGCGGATCCAAGCACATCCTGACCATCGAGGATCCGCCGGAATACCGCATTCGCGGCGAAGGCA
TCAACCAGACCCCACTGGTCTACGACGCCACCGACCCAGACGCAGAACGCCAGGCCTGGGCCGCGGGCATCGCCAACGGC
ATGCGCCTGGATCCGGACTACATGATGATCGGCGAAGTACGCGACCTCTTCGCCGCTGTCGCCGCGTTCCGTGGTGCGAT
GACCGGGCACGGCCTATGGTCGACCCTGCACACCAACAGCGCGATCGGCATTGTCCAGCGCCTGAAGGACCTGGGCGTCG
ACCCCGGCTTGCTGTTCGATCCGGCCCTGCTGACCGGCCTGATCAACCAGAGCCTGCTGCCCAAGCTCTGCCCCCACTGC
AAAGTGCGCTTCCAAGACGACCAAGACCAACTCGCGCCCGACTTGGTCGAACGGGTCCGACGCTTGACCGATGTTTCCCA
GGTTCACGTCAAGGGGCCTGGCTGCCAGGCCTGCCGTGGCTCCGGGGTCAACGGCCGCTCGATCGTCGCCGAGGTGGTTC
TGCCCACCCTCGCCTTCATGCGTGTGTTCGCCAAAGGCGGCCCAGCCGAGGCACGCAACTACTGGGTCAAGACCATGCAG
GGCATCACCAAGCACGCCCACGCCATCCGCCGCATCAACGAGGGCATGTTCGACCCGCAGATGGTCGAGGATTTCATTGG
GCCACTCGACTTCGATGAGCATCTGCTCGACGACAGCTTCTACTCGCAGGAGGCGTGCTGA:

Protein sequence: (SEQ ID NO: 183)

MTNLQIAALAQPSMVTQLLTADGGEWEVSKHLQEIMALAADGTLYLSESHQNDIHVLSFIDRLDRRGFRYQLNLTDLQTI
HQLYRAVAMDGLVDSDGQRATQMQERVVKIIRKATELRASDVHFVVSPAGTGSKIRFRVDGLLKTVEQFRSQELHELCAT
IYQSMCDVAEPLFKPQLDQDARMSQTFVEKLNLFSARIATRPRAGGFLMILRLLYDDTGLDSLEQLGYLPEQNALFDRMM
RMPYGINILSGPTGSGKSMTLKVTLEGLDKLHGGSKHILTIEDPPEYRIRGEGINQTPLVYDATDPDAERQAWAAGIANG
MRLDPDYMMIGEVRDLFAAVAAFRGAMTGHGLWSTLHTNSAIGIVQRLKDLGVDPGLLFDPALLTGLINQSLLPKLCPHC
KVRFQDHQDQLAPDLVERVRRLTDVSQVHVKGPGCQACRGSGVNGRSIVAEVVLPTLAFMRVFAKGGPAEARNYWVKTMQ
GITKHAHAIRRINEGMFDPQMVEDFIGPLDFDEHLLDDSFYSQEAC.

RL083
DNA sequence: (SEQ ID NO: 62)

ATGCGAACTGAGCCGATCGGCATGGCCGTGGCCGTGCTCTTCCTCCTCGCGTCTGGCCAGGCCTGCGCTGGCACCGTTGG
CGAACTTGCGGAGATCCAGGCCCAGGCCATCCTCACCGAGGCCAAGGTGCGCCTGGCCACGGCGCAGCGGCAATTGGAAG
GCAAAGGCGAAACCGGCCAGGTCGTCAGCGCCCAGGGGCAGACGTTCGCCATGCCGGTGCCGGCGGCGCCGCCGACGATC
ACGCAGCCGGTTCCGCCAGTGGTGCGGACCATCTACGGCGCCGGCGGCAAGATGACTGCCACGTTCTTGTTCCCGGGCGG
GTACGAGGTTGACGCCGCCAGCGGCGCGGAGCTGCCTGGCAAATACCGCGTCGAGTCAATCTCGCTGGACCAGGTCGTGC
TCACCGACAAGGACGGCAACCGCGTGCCCGTGGGCTTCTCCAGCGTTGCCCCCACCCAAGCCTCCTCTACGGCCCAAGGC
GCCTCGGTTCCGCCGGCGCTGCCCGGGGCTGTACCGCAGCCGTTCATTCAGTAG

Protein sequence: (SEQ ID NO: 184)

MRTEPIGMAVAVLFLLASGQACAGTVGELAEIQAQAILTEAKVRLATAQRQLEGKGETGQVVSAQGQTFAMPVPAAPPTI
TQPVPPVVRTIYGAGGKMTATFLFPGGYEVDAASGAELPGKYRVESISLDQVVLTDKDGNRVPVGFSSVAPTQASSTAQG
ASVPPALPGAVPQPFIQ.

RL084
DNA sequence: (SEQ ID NO: 63)

ATGGAGAAGCCTGACCTCGGCAGCCGTGGACCAGACGTCTCGATCCTGAGCTACCACGGCAACAAGTTCGTCAGCGGCCT
GTTCTGGCGGCCGCTGTCCAGCCAGCGGCAGTACATGAAGGAAGCGCGCAAGCTGGGCAAGGAAGAGCATCTGGACATCG
TTGCCATCCGCCATTCACCGACGGTGATCCAGGCCGGCTTCGTTTCGAAGTCGCAAGGCGCAGTCAAGGGGATGTACTCC
CTGGCCTCGGCGCTTTCAGGCCAGTTCGACGGCGACTTCCTGGCCTGCTGGAAAGTCGACGAGGACCGCTACGCGCTGGT
CGCCACGCTCGATGGCGCGATTGTCCCCGGGCAGGATCTGGTCACCACCCTCGACGAGGCCCGGGACCGGGTCAGGAAGC
TCTCTACGCGCGGCGTGCTGCGAAACGCACAGGTCTTCGTTCCCGAAGGGTTCGATTTCCCCGTCAAGGACTTCGACATC
GAGGAACTGCTCGCGCCGAAGCGCCTGCGGCGCGACTACCGCCTCCGGCAACTCACCTTCGGCTTGTCCGCCAGGGAGTG
GACGGCAGTGGCCCTGCTCGGTTGCGTGGTAGGTGGGTCGCTAACCGCCTACTACCTATGGAATGCCCACCAGGAAGAGC
TCGCCAGGCAAGCCGCGCTCCTCGAGGAGCAGAGGCGCCTCGCCGAGCTGGCCGAGAAGAACGCCCAGGCCAAGCAGCCG
CTGGACCTGGCGTCATTGCAGAAGCCTTGGACGCTCATACCTGACCTCGAGGACATGCTACGCGCCTGTAGCAAGGCAAC
GGGGGTACTGTCGCTGTCGATCCAGGGCTGGCTCTTCGAATCCAGCAAGTGCGACGGCAGGGTCCTGGTCGCCACCTACC
ACCGTACCGGCAACAGCACAGCAGCCGACCTGACAGCGGCCAGCCAGCACCTGTTCGCCGACCGCCCCGCCTTCGTCATC
GACAACGGCAACACCGCGGCCCTGAAGGTCGATCTGAAGGTTGCCATCGGCAGTGATGAGCCGCTACTGCCGGCGGACGA
CGTTCTGCAGGCGCTGACGAGCCACCTGTACCGTCAAGGGGTCGAGCCCAAGCTGTCGATCAGCCAGGAGACAACTCCGC
CCCTCCCTGGCGCGGAAGCTGCGACTGAACAGCAAGTGGTGTTGCCTTCCTGGAAGAAATTCACCTTCAGCGCCCAGACC
CGGCTCCCGGCAGACCTGACCTTCCAGGGGCTGCCCGCTGCCGGTGTCCGCATCACCAACCTCGAAACCACGCTCAAGGA
CAGCCAGTTGGACTGGACTGTCACAGGAGAAATCTATGCGAACTGA

FIGURE 30K

Protein sequence: (SEQ ID NO: 185)
MEKPDLGSRGPDVSILSYHGNKFVSGLFWRPLSSQRQYMKEARKLGKEEHLDIVAIRHSPTVIQAGFVSKSQGAVKGMYS
LASALSGQFDGDFLACWKVDEDRYALVATLDGAIVPGQDLVTTLDEARDRVRKLSTRGVLRNAQVFVPEGFDFPVKDFDI
EELLAPKRLRRDYRLRQLTFGLSAREWTAVALLGCVVGGSLTAYYLWNAHQEELARQAALLEEQRRLAELAEKNAQAKQP
LDLASLQKPWTLIPDLEDMLRACSKATGVLSLSIQGWLFESSKCDGRVLVATYHRTGNSTAADLTAASQHLFADRPAFVI
DNGNTAALKVDLKVAIGSDEPLLPADDVLQALTSHLYRQGVEPKLSISQETTPPLPGAEAATEQQVVLPSWKKFTFSAQT
RLPADLTFQGLPAAGVRITNLETTLKDSQLDWTVTGEIYAN.

RL085
DNA sequence: (SEQ ID NO: 64)

ATCGTGTGCGAAGCTACGGCAGATTCCGCGTCTACGATCGCAGCGCAGGTGCGCAACACCCGACCGGATCGGCGCGATAC
GGTGGTGTTCTCCGACAAACCCTGGGTCAGCACGAAACCCCTAAGCGTTTCGCACACCTTGTCCAGTGACTGCATCGTGA
CGTGGCGCCCTGCAGGCGCAGCGTCGCTGCAGGAGGCCGCCCAGGAAGTCATCAACCAATGCCACATGGCGGTCAGTATC
ACGCCCGACGCGCTGAACCCGGCCGCCTTCGCCGTGCAACCTCAGCAGCGCGCGAGCAACGCCCCGCCGCCCATCCAAGG
CGGCCAGGACATGGCCACCATGCTGTTTCCTGCCTCCGTCGCCAACGGCATGTCGCTCGGTGCCGGCGGCAGCATGGGGT
CGAGCTTCGGGTCCTACGGTCCGCGGTCTCTGTACAACATCAAATGGAACGGCAAAGTCAGCGGGTTCCTCGATCTCATC
GCCGCCCGAGCCGGCGTGTCCTGGCGCTACAACCCAACCGAGAAAAGGGTCGAGTTCTACTACCTGGACACTCGGACCTT
CCGCATGTACGCCTTCGACGACGTCAACACGGTGGACTCCACCGTGCGTTCCGGTATGACGACGGCCGCCGGCATCAGCG
GGGACGGCTCCGGATCCACCGGACAGAATGGCAGCTCCGGCATCAGCGGCGACTCCGGCAGCAAGCAGACCACCAGCTCG
GAGCTGAAGACATCGATCCTCAGCGACATCGAGAACAGCATCAACTCGATGCTGACGCCGAGCATGGGACGCATGTCGCT
GTCGCGTGCCACGGGCACCCTGACCGTCACCGACCGTCCAGAAGTCCTCAACCGTGTCCAGCAGTTGGTCAACCGAGAGA
ACGAGAGCATCACCAAGCAGGTGCTGCTGAACGTCAACGTGCTCTCGGTCGCCCTGACCGACAAGGATCAACTGGGGATC
GACTGGAACCTGGTCTACAAGTCGCTCAACAACAAGTGGGGCATCGGCCTGAAGAACACCATGCCGGGCATCGATCAAAG
CGCGATCTCCGGCTCCGTGAGCATCCTGGATACCGCCAACAGCGCCTGGGCAGGATCCAAGGCCATGGTCCAGGCGCTGG
CCCAGCAGGGCCGCGTCTCGACCGTCCGATCCCCGTCCGTGACCACGCTCAACCTCCAGTCGGCGCCGATCCAGATCGGC
CGCTACGACAGCTACCTGGCCTCCAGCCAGATCTCCAACGTCGCCCAGGTCGGCAGTACCACCTCGCTGATCCCGGGCGC
CGTGACCAGCGGCTACAACATGAGCCTGCTGCCGTTCGTGATGGAAAGCGGCGAGATGCTGCTGAAGATCAACATCAACA
TGACCTCCCGGCCGACGTTCGAAATGCAGACCAGCGGGGACTCCAAAGCCCAGTTCCCGAGCTACGACATACAACTGTTC
GACCAGAAGGTACGTCTGCGCAGCGGCGAGACCTTGGTACTCTCCGGCTTCGACCAGACCACCGAGGACACCAACAAGGT
CGGCACCGGCGACGCTGGGTTCTTCGGTCTTGGCGGCGGGCTGACCCGCAATACCAAGCGCGAGGTCATCGTGGTGCTGA
TCACCCCCGTCGTGCTGGGCTGA

Protein sequence: (SEQ ID NO: 186)
IVCEATADSASTIAAQVRNTRPDRRDTVVFSDKPWVSTKPLSVSHTLSSDCIVTWRPAGAASLQEAAQEVINQCHMAVSI
TPDALNPAAFAVQPQQRASNAPPPIQGGQDMATMLFPASVANGMSLGAGGSMGSSFGSYGPRSLYNIKWNGKVSGFLDLI
AARAGVSWRYNPTEKRVEFYYLDTRTFRMYAFDDVNTVDSTVRSGMTTAAGISGDGSGSTGQNGSSGISGDSGSKQTTSS
ELKTSILSDIENSINSMLTPSMGRMSLSRATGTLTVTDRPEVLNRVQQLVNRENESITKQVLLNVNVLSVALTDKDQLGI
DWNLVYKSLNNKWGIGLKNTMPGIDQSAISGSVSILDTANSAWAGSKAMVQALAQQGRVSTVRSPSVTTLNLQSAPIQIG
RYDSYLASSQISNVAQVGSTTSLIPGAVTSGYNMSLLPFVMESGEMLLKININMTSRPTFEMQTSGDSKAQFPSYDIQLF
DQKVRLRSGETLVLSGFDQTTEDTNKVGTGDAGFFGLGGGLTRNTKREVIVVLITPVVLG.

FIGURE 30L

RL086
DNA sequence: (SEQ ID NO: 65)

ATGACCAGGCAGTTGACCACTCTCACGCTGTGCCTGCTGCTCGCCAGCTGCACGACCCACAAGGCTGAGCCGGCCAGGCC
AGCCTTCGACAGCAGCCGCAATCCAGACCTGCTTTCTCCGGACCTGTATCCAAACGGTGTGCAGCCGGAGAAAGAGCCCG
TAGTGCGCTATGGGCGCTACACCCTGGTCAGCACCCAGCCTGATGCCGGTCAACGCGACCTGATGGCCCAGATCATCGAC
GTAACCATCCCGTCGAGCATGAACCCGAGCGTCAAGGACGCCATGCAGTACGTGATGAGCCGCTCGGGTTACTCGCTGTG
CCCGGCAGACGCCGGTCATGTGAACATCCTCTACACCCGGCCGCTGCCGGCAGCTCAGTACAAGCTCGGCCCGATGACCC
TGCGCAACACCCTCCAGGTCCTCTCCGGCCCAGCCTGGCAGGTTAAGGTCGACGAGGTCGCGCGGCAGGTCTGCTTCGTG
CTGCGCCCGGGCTATCAACTTCCCCCGGCGCCGAGGCCGAAACCGGTCCAGCAACTGTATGCGAAGCCCGCTGCCCCAAC
TCCGCCGGCGGTAGCGCAACCCTCCTCCACGGAGAAAGTCAGCACGCTGGAGTCGCCCATCGTGGTCGCCTCGGTGCCGA
CACCGGCGCCGATCACAACCAGCCACGCTCCGGCCAAGAAGCCTGAATCCACCACTGTGCTCCCCCCAGCCGCACCGGCC
AAGGATGGCCACCCCTCTTCTCCTCCCGCGGCTTCGGCACCGACCAAGCCTGCGGCCTCCGCCGTGAAGTCCACGCCGCC
CACTCCACCCACCGTGGCTTCCGCCCCACCGGTCAAGGTGCTCACGCCGCCGGAACCGAGCCGGCCGCTGGCACAGGCCT
GGTCAGCCGAGACGGGATCAACCCTGCGCGACACCTTGGAAGCTTGGGCAAAGCGCGCACGCTGGACCGTCCGCTGGGAG
CCGCAGGATCTCAACTATCCGATCGAGGCTCCACTGACCTTCCACGGCTCCTTCGAGGACGCGGTATCCGAGCTGTTCCC
CCTGTATGACGCTGCCGAACGGCCCTTCCTGGTGAACGCCAGCCGCCCGCAGTCCCTGATCATCATCAAGGAGCGCAAGA
ACTGA

Protein sequence: (SEQ ID NO: 187)
MTRQLTTLTLCLLLASCTTHKAEPARPAFDSSRNPDLLSPDLYPNGVQPEKEPVVRYGRYTLVSTQPDAGQRDLMAQIID
VTIPSSMNPSVKDAMQYVMSRSGYSLCPADAGHVNILYTRPLPAAQYKLGPMTLRNTLQVLSGPAWQVKVDEVARQVCFV
LRPGYQLPPAPRPKPVQQLYAKPAAPTPPAVAQPSSTEKVSTLESPIVVASVPTPAPITTSHAPAKKPESTTVLPPAAPA
KDGHPSSPPAASAPTKPAASAVKSTPPTPPTVASAPPVKVLTPPEPSRPLAQAWSAETGSTLRDTLEAWAKRARWTVRWE
PQDLNYPIEAPLTFHGSFEDAVSELFPLYDAAERPFLVNASRPQSLIIIKERKN.

RL087
DNA sequence: (SEQ ID NO: 66)
TTGAGCTTTAAATACTATTGGGCTAAATTTTTCTGGGGAGCTTTCTTCTTTGTTTTAGTCGCTTGGAAAGGCTCCGTATT
TCCAAGCCTGGCATCAGTTAACCCCTTGGTAGTGGCTGGATTCAGTACTATCCTGTTTCCTTTCTCGGTAAGGCTTGTTG
AAGACTTCGCTTTAAAATATACGGAAAAAGAGTTCTGGGTCACAGGTTTTTTCTCCGAAACCCCTGCAAAAACAGGATTG
TATGCAGTCTTTTATTTGGCTTGTTATTTGTTTTCAATTCCCTTGGGGATGATTTTTTTATTCTATAAATACGGAAAGGC
CTCGTAG

Protein sequence: (SEQ ID NO:188)
LSFKYYWAKFFWGAFFFVLVAWKGSVFPSLASVNPLVVAGFSTILFPFSVRLVEDFALKYTEKEFWVTGFFSETPAKTGL
YAVFYLACYLFSIPLGMIFLFYKYGKAS.

RL088
DNA sequence: (SEQ ID NO: 67)

ATGTCCAATGACAACGAAGTACCTGGTTCCATGGTTATTGTCGCACAAGGTCCAGACGATCAATACGCATACGAGGTTCC
CCCTATCGATAGCGCGGCCGTTGCCGGGAATATGTTTGGCGACTTGATTCAAAGAGACATATATCTACAGAAAAACATTT
ATTATCCAGTCCGATCCATTGTTGAACAAGGAACAAAAGAAAAGAAGGAGATCAACAAGAAAGTATCTGATCAAGTCGAT
GGCTTGCTAAAGCAGATCACTCAAGGAAAAAGGGAGGCCACAAGGCAAGAGCGAGTCGATGTCATGTCGGCAGTCCTGCA
CAAGATGGAATCTGATCTTGAAGGATACAAAAAGACCTTTACCAAAGGCCCATTCATTGACTACGAAAAGCAGTCAAGCC
TCTCCATCTATGAGGCCTGGGTCAAGATCTGGGAGAAGAACTCTTGGGAAGAAAGAAAGAAGTACCCTTTTCAGCAGCTT
GTTAGAGATGAACTGGAGCGGGCGGTTGCCTACTACAAACAAGATTCACTCTCTGAAGCGGTAAAAGTGCTAAGACAGGA
GCTCAACAAGCAAAAAGCGCTAAAGGAAAAAGAGGACCTCTCTCAACTGGAGCGGGACTACAAAACCAGAAAGGCGAATC
TCGAGATGAAAGTACAATCCGAGCTTGATCAAGCGGGAAGTGCTTTGCCTCCATTGGTCAGTCCAACGCCAGAGCAATGG
CTTGAACGTGCCACAAGACTGGTTACGCAAGCAATTGCTGATAAAAAGCAGCTGCAGACCACAAACAATACTCTTATCAA
GAATGCCCCAACCCCTCTAGAAAAGCAGAAAGCCATCTACAATGGTGAGCTACTTGTGGATGAGATAGCCAGTCTACAGA
CCCGCTTAGATAAGCTGAACGCCGAAACGACACGACGCAGGACAGAAGCAGAACGCAAGGCGGCCGAGGAACAAGCGTTG
CAAGATGCTGTTAAATTTACTGCCGACTTTTATAAGGAAGTAACTGAGAAATTTGGCGCACGAACATCAGAGATGGCGCA
CCAACTGGCCGAAGGCGCCAGGGGGAAAAATATCAGGAGTTCGGCGGAAGCAATCAATTCGTTTGAAAAACACAAGGATG
CGTTAAATAAAAAACTTAGCCTTAAAGATAGGCAAGCCATTGCCAAAGCCTTTGATTCTCTAGACAAGCAGATGATGGCG
AAGAGCCTTGAGAAATTTAGCAAAGGCTTTGGAGTTGTAGGCAAAGCTATTGACGCCGCCAGCCTGTACCAAGAGTTCAA
GATATCTACGGAAACCGGGGACTGGAAACCATTCTTTGTAAAAGTTGAAACACTAGCTGCTGGTGCGGCCGCCAGTTGGC
TTGTGGGTATTGCATTTGCCACGGCAACGGCCACTCCTATAGGCATCCTGGGGTTCGCACTGGTAATGGCAGTTACCGGG
GCGATGATTGACGAAGGCCTTCTAGAAAAAGCAAACAACCTTGTAATGTCCATTTAA

FIGURE 30M

Protein sequence: (SEQ ID NO:189)
MSNDNEVPGSMVIVAQGPDDQYAYEVPPIDSAAVAGNMFGDLIQRDIYLQKNIYYPVRSIVEQGTKEKKEINKKVSDQVD
GLLKQITQGKREATRQERVDVMSAVLHKMESDLEGYKKTFTKGPFIDYEKQSSLSIYEAWVKIWEKNSWEERKKYPFQQL
VRDELERAVAYYKQDSLSEAVKVLRQELNKQKALKEKEDLSQLERDYKTRKANLEMKVQSELDQAGSALPPLVSPTPEQW
LERATRLVTQAIADKKQLQTTNNTLIKNAPTPLEKQKAIYNGELLVDEIASLQTRLDKLNAETTRRRTEAERKAAEEQAL
QDAVKFTADFYKEVTEKFGARTSEMAHQLAEGARGKNIRSSAEAINSFEKHKDALNKKLSLKDRQAIAKAFDSLDKQMMA
KSLEKFSKGFGVVGKAIDAASLYQEFKISTETGDWKPFFVKVETLAAGAAASWLVGIAFATATATPIGILGFALVMAVTG
AMIDEGLLEKANNLVMSI.

RL089
DNA sequence: SEQ ID NO: 68

ATGAACCGTCCACGCCTGGTTAATCGTACCTCCGCGACACCTTCGACGCTTCTGCAGCGGGCTATCTTCGACGGCTACGA
CTTCGGCTTGAAGATCCCCTACATCGCAGGCAGCAATCGCGCGCTGCTGGAGCTGTCCGGCTTCTTCATCAGCGCCCGGG
AGCATCCGTTGCACCGCTACTGGCGGGTCCCCAAAGGCAAGCTGCTGCCTGAACTGGACACTCTGTACAACCGTCTCGCC
GAGCTAGCTGGAGGCCTTCACTCCCAGTCCTGGCGGGAGTTCAGCTCCTTGGTCGAATCCGCGCAGGCCTCGCTTGACCG
ACAGGCCTTCACCTGGGGGATGCTGCTGCGCATCGCGCCCCTGGCCGAGGGCGGCGTCCTACTGTCAGGCGAGTTCCATC
CTGGTGTTGTGGCGGTGGCTCGGCGGATGCGCGGGGTATTCCTGCGCCCATCGAGTTCCTGGCGCATCGACACCACTCCC
GAGCTGCTCCGAAGCAACCTGATTCTGGAGCTTGGCCTCGCCGAGGAACAATTCGAGATTCTGGATACTGTCCAGGAGCT
GCTCAGCGACGGCAGCTTCGCGCCGTCGACCGAGCTGCCCAGCATGAGCATCGGCGGTCCACAGCAGGAACCGGCAGCGC
CATCCCTGGAGGACGAGTCAGCCTCTGACATCTACCTCGCCGCGGTGCCGGAGATCGAGCGCACCGAGTACAGCTCGGCT
GATATCGAGGCGGCGCTTCAGGGCTACTCTCTACTGGCCCACCAGCCTGACGGCATCGCTCATCTGCTGCAGAGAACCAG
CGCCTTATTGGCCGACGACATGGGATTGGGCAAGACCCGCCAGGCGGTCATCGCCGCTTCGATCCGCGCGGCGGGCAGAC
CAATCCTGGTCATCACCCTGGCTACCCTGCTGATCAATTGGCAGCGGGAGATCCAGGAGGTCTATCCCTCGGCCACCGTG
GCCATCCAGCAGGACACCCCAGAGGCGCAGTGGATCCTCGTCAACTACGAGCAGTTGAGCCCCTTCGTCGCCAACGCTTC
GCGCTTCGCCGTGATGGTCATCGACGAGGCGCAGCGGATGAAGGAACCGACGGCGCAATGCACGCGGCACGGTTTCGACA
TTGCCGCCCAAGTGCCGAACCGCTACCTGCTTACCGGCACGCCGGTGCTCAACCGCGAGACAGAGCTGCACACCCTGCTG
CGCCTCTCAGGCCACCCCATCGGCCAACTGCCGCTGAAAGAGTTCTGCGACCGTTTCGCCGGCAACCCGGAGTTCCGCCA
GAGTCTGCGGGCGGAGCTGGGTGACTGGATGCTGCGCAGGCGCAAAGATGTGCTGCCCAGCCTCAAGGGCAAGCAGCGGC
AGTTGCTGAAGGTGGCCCTCTCCACCGAGGAACGCCAGCAATACGACGTGCTGCGCCTCGAGGACCGACCGGTCTTCGCG
CGACTCGGCGCGCTGCGGCGTTACCTGGAAACGGTGAAAGTTCGCGTGGCGATGGACCTGTTGAGCGAGCTCGACGCAGA
GGACAAGGTGATCCTGTTCTGCGAGTTCAAGCCGACCGTGGCTGCGCTGAAGGAACTCTGCGAGCAGGCCGGACACGGCT
GCGTCACGCTGGTGGGCAATGACTCGCTCACCAAGCGGCAGAAGGCGATAGATCGCTTCCAGCAGGATCCCGACTGCCGA
GTGTTCATCTGCACTACGGCGGCCGCAGGGACGGGCAACAACCTCACTGCGGCGAACTACGTGTTTTTCCTCGGCCTGCC
CTGGACTCCCGGTCAGCAGGAACAAGCCGAAGACCGCGCGTACCGAAACGGCCAGCTCCGCATGGTCGTGGTGAAAATCC
CACTGGTCGAGGCCACGATCGACGAGCAACTGTGGCAACTGCTCAACGCGAAACGCCAGGTTGCCCAGGACCTCATCGAG
CCCGAGCAGGTCGACGGAAACCGCGCGCTTTTAGCCGCAAGCCTAACTGGATAA

Protein sequence: (SEQ ID NO: 190)
MNRPRLVNRTSATPSTLLQRAIFDGYDFGLKIPYIAGSNRALLELSGFFISAREHPLHRYWRVPKGKLLPELDTLYNRLA
ELAGGLHSQSWREFSSLVESAQASLDRQAFTWGMLLRIAPLAEGGVLLSGEFHPGVVAVARRMRGVFLRPSSSWRIDTTP
ELLRSNLILELGLAEEQFEILDTVQELLSDGSFAPSTELPSMSIGGPQQEPAAPSLEDESASDIYLAAVPEIERTEYSSA
DIEAALQGYSLLAHQPDGIAHLLQRTSALLADDMGLGKTRQAVIAASIRAAGRPILVITLATLLINWQREIQEVYPSATV
AIQQDTPEAQWILVNYEQLSPFVANASRFAVMVIDEAQRMKEPTAQCTRHGFDIAAQVPNRYLLTGTPVLNRETELHTLL
RLSGHPIGQLPLKEFCDRFAGNPEFRQSLRAELGDWMLRRRKDVLPSLKGKQRQLLKVALSTEERQQYDVLRLEDRPVFA
RLGALRRYLETVKVRVAMDLLSELDAEDKVILFCEFKPTVAALKELCEQAGHGCVTLVGNDSLTKRQKAIDRFQQDPDCR
VFICTTAAAGTGNNLTAANYVFFLGLPWTPGQQEQAEDRAYRNGQLRMVVVKIPLVEATIDEQLWQLLNAKRQVAQDLIE
PEQVDGNRALLAASLTG.

FIGURE 30N

RL090
DNA sequence: (SEQ ID NO: 69)

GTGGCACCTCTCGACAACGCCCCCCCTAGCGGGCCGCTACAGGATCCATCCCTGGCCCGCTACAGCGAGCGGCAGCTCGC
CGTCGCCAACACCTGGGCAACACATTTCTCCCTCGCAGGGACAGCTCGAACCAAGTTCATCCGTCACTACCTGCGCAGCA
CATCTACGACCAGGTGCTGGTGCATCACAGTCGCTGCTGACAACGGTGTGCGTTACACCATCATGCGTGCAGGGCCGCTA
CTCCAGGTATTCGACGGTCAACTAATTGGTGCGTGGGAGTGCAAGCCTGCCCATCGTATCCCGGCAAGCACGCCGTCTCG
AGCAGGGGCCTTGAAGCTGCTACAGCGCCTTCAAAAGTTCGACGACGCAGTTGCTGTACTCAGCTCATACACAAAGCGAG
CGCACGACCTAGCCACACAGATGGCCAGGGACGATCTCGGACTTCAACATCGCCTCGTGTATCCGAGCCACAGCAACAAG
CGCTACTACGCGCCAAGGCACCAGTTCTACTTGAAGCAGATCGGAGCGGTCTTGCGAACCTTCAGACAGGTCCTGGACCA
AGACCTGCTGTTCGCCATCCGCTCGGTTCGGTGCCTCTCGCCCCAGCTCTACAACTGGCTGGCTCAAGGCGACCAGGTGC
GCCGGCTGCAAATGCTGAAGGCTCAGCCGGTCTTGACGCCGCTACTGGTGGATTGCGAGGAGGGAGTCTGGCCTCACACG
ACGACCAACGACAACGGCGAGAGCATCCGCCATTACCTTCCTTGCCCCTTTCCCCAGCTTGACAGTGAACGACCGCAGGC
CGCCGCCATGCCATGCGACTTGTACCTCGATATGGGCCGTATTCTTGGGCAGGTCGCGGACGAAGGAATTTCGGTCATCA
ACTTTTTCGCCTGGCTATTTCAGGCGCCGCGGGCCTCGATTCGATTTCTTAGTCACGTCAGTCCCGGCCGTGCGGGAGGA
GCTCTCTTCCATCGCAAACGGGAAGGCCGACATTCGGGATGGCATGCTCTCCTACTGGCGGCATCGCTAGGTAACCGGCG
GCCGATCACTCGCGCTCAATGGACAGCATTCTATGCCGCCTACAATGCGATCCCTTGGCAAGTTCACAACGCCAAGCCCG
ACTACAACCGTCTCTTCAACGGCTGCCCGTCGGATTGGCAGGATCCGGCATGGCTTGCAATCACTGCACGGCTGAGAGAC
ATCAAGGAGTTCTATACCGCCCTCGACCAGGGGAACTCACAGGTTGTTCGGCAGGCGCGCAGCGCCCTGAAAGCGTATCT
GGGTCATTGTACCTACCGACAAGCTGGCAACCTGGTGGACGACTACCACCAGGTCCAGAGGGAGCTGCGTGCCGCAGTGC
AGAGCAGCCTGCCCGATCTGGTCGACACCGACGAGTACACCACCTGGGAGGGAATGCTGTCTGTCGGTCTTATCGATTGC
CCTAATGGACTGCAGATCGTCGAGCTCCGCTGTCCTGCCGACCTATATGCCGAACATATCGCTCTGGCACATTGCATCGA
TAGCTACGACCAGGCCGCCTACCGAGGAGACTGCCGACTGCTCTCAGTACGTGAGGCTGGTCGTCCGCTGGCCTCTGCCG
AATTGGAGCTCAGGCGTGAGCATGGCGAGCCTATAGGTAGGCCCTGGAGTCCCAAGCACCTTTCCACGGTGCAACTGCGC
GAATTCGATAATGCCCCCGTGCCGACCGACTCGCCTGCCGGCCAGGCATACCGCTGGTTCATGGAACGAATTCGCTCTGG
AGCCATAGCGACGAACCTGAACTGGCCCGACATGACCGTCCACATGACGCGCTTCGCCAATGGTCGCTGGAAGGCGGGCC
TCGCCGAAGCCACGGCGAAGTGGCTGCTCACTCAGTTGGAAGACCGATGA

Protein sequence: (SEQ ID NO: 191)
VAPLDNAPPSGPLQDPSLARYSERQLAVANTWATHFSLAGTARTKFIRHYLRSTSTTRCWCITVAADNGVRYTIMRAGPL
LQVFDGQLIGAWECKPAHRIPASTPSRAGALKLLQRLQKFDDAVAVLSSYTKRAHDLATQMARDDLGLQHRLVYPSHSNK
RYYAPRHQFYLKQIGAVLRTFRQVLDQDLLFAIRSVRCLSPQLYNWLAQGDQVRRLQMLKAQPVLTPLLVDCEEGVWPHT
TTNDNGESIRHYLPCPFPQLDSERPQAAAMPCDLYLDMGRILGQVADEGISVINFFAWLFQAPRASIRFLSHVSPGRAGG
ALFHRKREGRHSGWHALLLAASLGNRRPITRAQWTAFYAAYNAIPWQVHNAKPDYNRLFNGCPSDWQDPAWLAITARLRD
IKEFYTALDQGNSQVVRQARSALKAYLGHCTYRQAGNLVDDYHQVQRELRAAVQSSLPDLVDTDEYTTWEGMLSVGLIDC
PNGLQIVELRCPADLYAEHIALAHCIDSYDQAAYRGDCRLLSVREAGRPLASAELELRREHGEPIGRPWSPKHLSTVQLR
EFDNAPVPTDSPAGQAYRWFMERIRSGAIATNLNWPDMTVHMTRFANGRWKAGLAEATAKWLLTQLEDR.

RL091
DNA sequence: (SEQ ID NO: 70)

ATGCGAAAAGAGAATATATCTGCCGAAATCACAGAGCGAGCTTTTGATTTTTTCTATTGGTTCTCGCGATTTGAGTTCAG
CCTCAAAGAGAATGGCTACTTAAAAAATTACAAACCTGGAGCTAGGGCAGAGCCGGGATGGGAAAATTTTGTACAAAACC
ATTCTGACAAATACTCTCTTTCCCAATCAGCCACAGCACTAATCGAGCAGAGTCCAGAGCAACAAATAGTCCTGCCCGGT
AGAGAGCTGGGTTGGCGTCCGGTTAAATTAGATGAGGACAAAAGCGACTTAGCTAGAGTCGCTCGCTTACTTAAGACCGT
GCGAAACAATCTATTTCACGGAGGCAAGCATGGTGGTGCCAACTGGGACAACCCAGCGAGGACAATACATCTTATTCTTT
TAAGTAAAGCTATCCTTGACGAGTTTGCTGCACTAGGAGACTTTGAGGCTGACTACAAGAGAATTTACTGA

Protein sequence: (SEQ ID NO: 192)
MRKENISAEITERAFDFFYWFSRFEFSLKENGYLKNYKPGARAEPGWENFVQNHSDKYSLSQSATALIEQSPEQQIVLPG
RELGWRPVKLDEDKSDLARVARLLKTVRNNLFHGGKHGGANWDNPARTIHLILLSKAILDEFAALGDFEADYKRIY.

FIGURE 30 O

RL092
DNA sequence: (SEQ ID NO: 71)

ATGCACATCGTAATCATTGAAGCCCCGGGCAAGCTGAAAAAGCTGAGGTCCCTTCTCCCCTCGATTCGTCCCGACGTGAC
CTGGCAGGTCGAGGCGACAGCCGGCCACATCAGAGACCTACCCGTTCACGGGCAGGATCCGCAGATGCTCACCGTCGGCG
TGGGCCAGGATTTCAAACCGCACTACCAGATCCTCTCGGGCAAGGAAAAAACCGTCGCACGGCTGAAGGAGCTGCGGCAG
AAAGCCGTGGAAATCTACGTCGCATCGGACCCGGATCGCGAAGGCGAAAGCATTGGCTGGCACATCCTCCAAGCTGCCGG
GATCAAGAACTACAAGCGCGTTGCCTTCAAAGAAATCACAAAGTCATGCATCACCGCCGAACTCAGCTCGCCGCGTCGCC
TGGACCTCCCGAAGGTCGCCTCGCAGGAATGCCGTCGCGTCATCGATCGCCTGGTGGGGTATCTGGTCACGCCAGAGTTG
CGGCGCGTGATGGGTAGGCCGACCACCGCCGGGCGCGTGCAGTCCGTCGCGGTGTACCTGGTGGTCCTGCGAGAGCGGGA
GATCCGCGCCTTCACAGCAATCAAGCACTTCGGGGTGGAACTGACCTTCGTTTCGCCCAGCGACGGCCGTACCTGGACGG
CGGAATGGGATCCAGTGCCCGTGTTTGCCAGCGAGGAGTTCCCGTATGTCCAGGATCGTCAACTCGCAGAACTGGTGGGG
GCTATACGTAATGTCATCGTCGAGACCTGCATTGATAGCGAAGAAACCGATGCGCCTCCGGCACCGTTCATCTCCTCCTC
GCTCCAGATGGCCGCCGGGAATGCGCTGAAGTGGTCACCCGACAAGACGATGAAGGTCGCCCAGCGGCTGTATGAACAGG
GGCTCATCACCTACCACCGGACGGACAACCCCAATATCTCGAAGGACTCGATGCCGGATATCCGTGCTGTCGCCAAAGCC
TTGGGGCTGAAGTGTGTTGAGCAACAGCGGATGTTCAAAGCGGACCAAGACGCCCAGGAAGGCCACCCCGCCATCACCCC
TACCGACTGGATGGCCGCTGCCGCCGGTGAAACTGCTGATGAGCAGGCGCTGTACCAGCTCATTCGAGTCCGCGCGCTTG
CCAGTCAGATCGAAGCTGCCGTGTACGCAGTGAGAACCATCACCCTCCTGGGCGTCGGCCCCGACAAAAAGCCGCTGCGC
TTCGGCGCCAAAGGGAAGCTGTTGAACGTGCCTGGCTGGAGAAAACTGCTGCAGGGTGATGACGCCGAGGAGCAGAAGAA
CGAAACGCCTTCAAACCCCATCCCGATCCCGGCGCTGGAGCCACGCCAGATACTCAAGGTCTACAGCGGCGAGGTCCTGG
AGAAGAAAACCACCCCTCCCAAGCGATTCACCGACGCCAGCCTGGTGGGCGAGATGAAGCGCCGCGGGATTGGTCGGCCA
TCCTCCTACGCCTCGATCGTGAAGAACATCATCGACAAGGGCCAGGTGCAGATGAAGGGGCGAAGCCTGATCCCCGGCGA
GCTGGGAGAGGCCACCATCGCGCTCCTGGAGCACAACTTCAGCTTCCTCAGCCTCGACTTCACCCGCAACCTCGAGGTCG
CCTTGGACCGGATCGCCAACAGCGAGGACACCTACATGAACGTGGTCCAGCAGTTCTACCAGCTACTACAGACAGAGCTG
CAGACACTCCGCGCGCTCCCCAGCGCACAGGACGAACCACGCGCAAGCTCCACCGCCAGTATCTCCTCGGCGCCGACCAG
CGACTTCCTTTGCGGCAAGTGCGGTCTGCCCCTGGTTCACCGCAAGAAAGCCGGCAAAGGCGGCTTCGACTTCTGGGGTT
GCAGCGGCTATCGAACAACAGGGTGCAAGGTTAGCTACCCCACCAAGAGCGGCCGGCCTGACTTCGACAACCCGCGCGGG
CTATAG

Protein sequence: (SEQ ID NO: 193)
MHIVIIEAPGKLKKLRSLLPSIRPDVTWQVEATAGHIRDLPVHGQDPQMLTVGVGQDFKPHYQILSGKEKTVARLKELRQ
KAVEIYVASDPDREGESIGWHILQAAGIKNYKRVAFKEITKSCITAELSSPRRLDLPKVASQECRRVIDRLVGYLVTPEL
RRVMGRPTTAGRVQSVAVYLVVLRREREIRAFTAIKHFGVELTFVSPSDGRTWTAEWDPVPVFASEEFPYVQDRQLAELVG
AIRNVIVETCIDSEETDAPPAPFISSSLQMAAGNALKWSPDKTMKVAQRLYEQGLITYHRTDNPNISKDSMPDIRAVAKA
LGLKCVEQQRMFKADQDAQEGHPAITPTDWMAAAAGETADEQALYQLIRVRALASQIEAAVYAVRTITLLGVGPDKKPLR
FGAKGKLLNVPGWRKLLQGDDAEEQKNETPSNPIPIPALEPRQILKVYSGEVLEKKTTPPKRFTDASLVGEMKRRGIGRP
SSYASIVKNIIDKGQVQMKGRSLIPGELGEATIALLEHNFSFLSLDFTRNLEVALDRIANSEDTYMNVVQQFYQLLQTEL
QTLRALPSAQDEPRASSTASISSAPTSDFLCGKCGLPLVHRKKAGKGGFDFWGCSGYRTTGCKVSYPTKSGRPDFDNPRG
L.

RL093
DNA sequence: SEQ ID NO: 72

ATGGATCAAAGCCTTTGCACATGCATGCCAACGCCAATCGTCAACCCCAAGGAGCTGCGACTGTGCCACATGTTAGTCGG
TAGAACTTTCCCGATAACATTGATCGCAGGCGACCATTGGTTGAGCTATGACGGCAGCGCCTGGTGGGTCGATGCGGATG
AGCCCGCGACGGAGGACGAGGTGGCGGCTCTGTTGGTCAAGGCTGGTGGTGTCACTACGTGCTGGTGCGGATAG

Protein sequence: (SEQ ID NO: 194)
MDQSLCTCMPTPIVNPKELRLCHMLVGRTFPITLIAGDHWLSYDGSAWWVDADEPATEDEVAALLVKAGGVTTCWCG.

RL094
DNA sequence: SEQ ID NO: 73

GTGGCAAGGGCTTCCGAATCGGAAATCTCGACCAGTACGAGGTGCAGTGTGTCAAAGAGAGCGACCGATACCGACAAGCT
GGACAGACGACACTTCAACGATCCCCACCGGACTGTACGGGCTATTGGTGCTGAGGCCGCGCGGAAAGGGCTACGGGTGT
TCGACTGCCCCTACAGTCATCCTGCGATGCGGGCGTCCTGGTTGAAAGGGTTTGCCCAGGAGCAGCAACAGCAGCTCGAC
TTCTGA

Protein sequence: (SEQ ID NO: 195)
VARASESEISTSTRCSVSKRATDTDKLDRRHFNDPHRTVRAIGAEAARKGLRVFDCPYSHPAMRASWLKGFAQEQQQQLD
F

FIGURE 30P

RL095
DNA sequence: (SEQ ID NO: 74)

ATGGCTACCCCCGTCTTCTGGGAAGCCAACATTGGCTCGGCGCCGGAGCACCGCAGCTTCCCCAACGGCAACAATCCCCC
GCGGCAGTTGCTGCGACTGAACGTGATGTTCGACAACTCGATTCCCGATGGCCAAGGTGGCTACAAGGATCGCGGCGGCT
TCTGGTGCAGCGTCGAATGGTGGCATCAGGATGCCCAGCGCTTCGCCGAACTGTTCACGAAAGGTATGCGCGTCAAGGTC
GAAGGCAGGGCCATTATGGACCGCTGGCCGGACAAAGAGTCAGGCGAAGAAGTCCAGGCGCTGAAGGTCGAAGCCTCGCG
CATTTCCATCCTTCCGCATCGCCTGGCCGAGGTCACCCTGTTGCCAACCCAGCATCAACAGTCTCGGAACGTCCCGCAGC
AACCTGCTCAGCAAGATGCGCAATCGCAGCAGGACTACGACAGCGCCTTCGACGACGACATCCCCATGTA

Protein sequence: (SEQ ID NO: 196)
MATPVFWEANIGSAPEHRSFPNGNNPPRQLLRLNVMFDNSIPDGQGGYKDRGGFWCSVEWWHQDAQRFAELFTKGMRVKV
EGRAIMDRWPDKESGEEVQALKVEASRISILPHRLAEVTLLPTQHQQSRNVPQQPAQQDAQSQQDYDSAFDDDIPM

RL096
DNA sequence (SEQ ID NO: 75)

ATGCGGCAGCTCGATAAGGACCAGCAAGGCGCTCTGGAACAAAGTGCCTTCCGCCCACTGCAACAAACTGCCTTCCAGGC
GCTGCAACACAGTGCCTCACTAAAAGGCCTTTTAAAGCCTTTTAAAGGTAATAGGGAGCTGGCCCAGTTGGCGGAACAGT
GCGAAGCCATGGAGCAGGGATTGCTTGAACTTGCCCAGGGACTGCTGGCCCAGGTTCGTCGCCCACCCTTCACTCTACTG
CCCACCCGACTCATCGAGCAGCGCACATCCGCCCGCACAACCTTTCTCCGCTGGCAGCACATTGCATCCCGGCGGATGGG
CGTCGGGGTGTGGACGGAAATGCTGCGCCAGGACAAGACCCCGGAATACCTGCTGCAAGACCTCTACGAGATGGAGCTGC
AGCGCATCACCCTCAACATGCAGATCAGCCTGATCCACTCCATCGGCAAGCAGGCCGCCGAGTGCGCGGAAAAGATGGGC
CAGGCCGAGGCCGAGTTCATGGGCCGACTGCAGCAGAGCACCAACCACCACTGA

Protein sequence: (SEQ ID NO: 197)
MRQLDKDQQGALEQSAFRPLQQTAFQALQHSASLKGLLKPFKGNRELAQLAEQCEAMEQGLLELAQGLLAQVRRPPFTLL
PTRLIEQRTSARTTFLRWQHIASRRMGVGVWTEMLRQDKTPEYLLQDLYEMELQRITLNMQISLIHSIGKQAAECAEKMG
QAEAEFMGRLQQSTNHH

RL097
DNA sequence: (SEQ ID NO: 76)

ATGGCTGAAACCCATCGGCTTCAGATCGGCTCTCTCCGCAGCGATGTCGCTCTGACGCTTCACACCTATCACGCCGCCCG
CATCTGGACCGGCCGGCAGAAGTCGGATGCCAAGCACAGCATCCTCGGCCTCTCCGGCTTCTGCGCATACGTGAATCGCA
TGCACCGCGGGGCAGCACAGGACGATCCGTACTCCGACTGGTGGCTGGTTCAGATCGAAGAGAAAGTCGAGAGCTGCCAA
GCCGCACTCGAGGCCATCGACCAGCGACTGGATGACGTCATGGCCAAGCTGCCCGCGACCCTCGATATCTCCGAGAACCT
GTCCGTTACACCGGTCAAGGTCCCGTTGTTCATCTCCAACCCTCTCGGCTTCAAGGCAGTCTATCTCCTGACCAACTACG
ACGAACTCGCCCGTCGAATCCTGCTGGCCCAGCACGTCGGGCTGGTCGGTCGCCGCGACATGGAGGTCTGGCTCGACGAA
GGTGCGTCGGTGCTGCGAAGCCTGTTTGGTCTGGCCCAGAGCTACCAGTTCTCGGGCGCCACTCGCGACGACTTCGCCGC
AAACAATGCTCGCGCCGAAGCCGCGCGGAAGATGTACGAGAAGTTCGGCGAGATCCCGCAGGACATCCTGGAGGGCACTC
GACGCTCGAACTTCGCTCCGCCGATCACCCGGGGCCGCTCTGACGGTGATGCCGATGATGACGCTGACCGTGTCGAACTC
GAGGACTGA

Protein sequence: (SEQ ID NO: 198)
MAETHRLQIGSLRSDVALTLHTYHAARIWTGRQKSDAKHSILGLSGFCAYVNRMHRGAAQDDPYSDWWLVQIEEKVESCQ
AALEAIDQRLDDVMAKLPATLDISENLSVTPVKVPLFISNPLGFKAVYLLTNYDELARRILLAQHVGLVGRRDMEVWLDE
GASVLRSLFGLAQSYQFSGATRDDFAANNARAEAARKMYEKFGEIPQDILEGTRRSNFAPPITRGRSDGDADDDADRVEL
ED

RL098
DNA sequence: (SEQ ID NO: 77)

ATGTTCTTGAGCATGGCCCCTTTCTTTTTGGTCGTTCTCGTTCTTTCTGCACTTTTTACAGATGCGTGGAACGACCGAGA
ACTCAGGCTGTTGTTAATGCTGATCGTGTTCGGGTATTCAGTAACCGTGTTGACCATTACGGTTGAGATGTATCGCTTTG
AAATGGCGGAAAAAGCGATGTGGGGAGCTTTATGCAACAAAGCCAACTACATGAACTGCCAACCAGATTACCAACGGTAG

Protein sequence: (SEQ ID NO: 199)
MFLSMAPFFLVVLVLSALFTDAWNDRELRLLLMLIVFGYSVTVLTITVEMYRFEMAEKAMWGALCNKANYMNCQPDYQR.

RL099
DNA sequence: (SEQ ID NO: 78)

ATGAGAAAGTCTCGGTCGGGCGTCGTGTTTTTTGGTGATGCGGCCCGCATCACTCTCCCAGGTCCTGACCTCCGCGCCGC
CGGCGAGCTGGGTGATTCCACTGGAATCACTCCACCAGGAGCCGACCTCCGCGCCGCCGGCGAGCTGGGTGATTCCACTG
GAATCACTCTGCCAGGGATCCACTTCGGTATCGGCGGCAAGATGGGTGTTTCGGGCCGAAACACTTCGCCAAAGCGAGGC
ATCACCACTCACGAGGAACTCAAACAATGTTCTTGA

FIGURE 30Q

Protein sequence: (SEQ ID NO:200)
MRKSRSGVVFFGDAARITLPGPDLRAAGELGDSTGITPPGADLRAAGELGDSTGITLPGIHFGIGGKMGVSGRNTSPKRG
ITTHEELKQCS.
FIGURE

RL100
DNA sequence(SEQ ID NO: 79)

ATGAGGCTGTCGCGCTTTCCCATTTCGACACTGCTGGACTCGGCCTCGGGGCATCTCGAGGCCCATTTGTATAAGAAGCG
GCTTGCTGCCGAAAGCGGCGAACCGCTGGCTCAACAATATTCCGGCATCATTTTCAGCGGCAATCCTCATGAAACCGTTC
CACGGCGCCTCCTCCTGGATAAGCGTCTTACTCCGCTGGAGCGGAACTGCTGGCAAGTTTTCCGCTTGCTCATCAACGAC
GACGGACTGACCGCGTTCCCGACATACGAGCAACTGCGCCCCTATCTCGGTATGCAGCCGGGCAAGATCGCCTCGCGCGA
AACCATCGCCAAGGCACTCACGGTCCTTCGTCTGACCCGCTGGCTCAGCCTCGGCCGACGCCTGCGCAACGACCTCAACG
GACAGGTCCAGGGCAACGTTTACATCCTTCACGACGAGCCTGTCTCTCCAGCCGAAGCCTTGGAGCTGGACACCGACTAC
ATGCAGTTGCTGAGCCAATCCACCGGTCACGGCAACCGAGCCATACGCGAAATCGGGCAGATCATCTGGCGGGAGTTCAG
GGATGATCCGGACGTGGGTCGCCGCCTCCCTACCCATCTGGAGAAGCTCGAGGGACGCTTGAACCACCAGCAATGGGCTA
TCGATAGTCAGCTCGAAGCGGATCCAGCGGCAGAGTTCGGCATCCGAACTCTGTCGGATTTACCTCATTCCACCCCGAGT
TCGGATGCCGAACTCAGTGAAATCAGCGGCAAGCAATGCGCTCTACCGCTGAGTTCGGATACCGAACCCCGACAGAATCC
GCCGAGTACGCCCTTGGTTCGGATGCCGAACTCATATAGTACGTATACATACAAACAAGATTCTGTATGTAAAAAGCCAG
TACAACCGCGAGCACGCGAGGAAGCCCATCCGAACTGGCAGGATCTCCTGCACGCACTGGAGGCCGAGCAACGGATCCAG
GCAGTAAGCGCGCTCAGACGGGTGTCCGAGGATCTTCGGCTACCCATCATCGAGCAGTGGCAGCACCGTTGTGCCGGCGG
AACAGTCAGCAATCCGTTCGGCTACCTCATGACGCTCATCCAGCGTGCAGTCCAGGGCAAGTTCAACGCTTCTTGGGCTC
CGGAAGAACCGGCTGAGCGAACCATCCCCGCAACGGAACGCCCCATTCGTGCTCCGGCACCATCAAGCCCCATAGCGCCT
ACACAGCCTCAGGTCCAGCCCCGGGGGGATACCCGGACAGGGAGCGAGGTCCTCAGCCGGCTCAAGGACCTCATTCGGCC
CAGGCACGGATCGAGCGTGCCATCCGAGCGGGGTGATGATTCATGA

Protein sequence: (SEQ ID NO: 201)
MRLSRFPISTLLDSASGHLEAHLYKKRLAAESGEPLAQQYSGIIFSGNPHETVPRRLLLDKRLTPLERNCWQVFRLLIND
DGLTAFPTYEQLRPYLGMQPGKIASRETIAKALTVLRLTRWLSLGRRLRNDLNGQVQGNVYILHDEPVSPAEALELDTDY
MQLLSQSTGHGNRAIREIGQIIWREFRDDPDVGRRLPTHLEKLEGRLNHQQWAIDSQLEADPAAEFGIRTLSDLPHSTPS
SDAELSEISGKQCALPLSSDTEPRQNPPSTPLVRMPNSYSTYTYKQDSVCKKPVQPRAREEAHPNWQDLLHALEAEQRIQ
AVSALRRVSEDLRLPIIEQWQHRCAGGTVSNPFGYLMTLIQRAVQGKFNASWAPEEPAERTIPATERPIRAPAPSSPIAP
TQPQVQPRGDTRTGSEVLSRLKDLIRPRHGSSVPSERGDDS.

RL101
DNA sequence: (SEQ ID NO: 80)
ATGTCGAAGTCGACGATCAATGAAGCGGTCCTGACGCAGGTGCTCAACCACCTGCGCAACGGCCAGCTCAGGCGTTGTGC
CGAGATGGGGCTGCGGCCGGAGATTCTGGCTCAGCTCCAACAGCCTGCCGTCATGAGCATCCTGACCAATACCCCGGTTT
CCTGGGTAGATGTCAGAGTGAACATCGACGTCATGGAGAAAATCCTGGCCACAGCCGAGCGCAGCGCGCAGGAAGACCTG
CAGATCGAACGCGCACTGAAGCTGGGAGCCACCACAACGATGATCCAGAGCTTTTTCGGTCTGTCGCCGGAGGACACCGC
CACCAAGCGCTTGATGCTGGAGATCCACCCGCGCCGCGGTCGCTGGCGGCAGCTCGATGAACAGATCGAGCGCCAGATAT
GGTTCCGCTGGGAGCACCTGATGCAGGAAAATCAGGTCCGCCTTGAAGACAGCATGGAGTTGCTGGACATCGCGATGATC
CTCACAGAGGAAATCAACGCCGGAATCGAACAAGACAGTCCAGAATTCATCAGCCTCGCCATTGTTTGGTCTCTCATCCA
GAGCTGGTTGAAAGACGGGCTCTATCCGTCTGGCAAATCGAGCCAGAGCCAGGCGGGCCTGCAAAAGTCCCAATCCACTC
TTTACCTCGCTAGCGTCAGCTCACACCTGCCCCACTCTGCCCCATCCGCAACAACGCAGGTGAACGCTGAGACAGAACGT
CAACAACTACTGAACCTGGTTCAGTCGGAAGGCGACACAGCACCATGA

Protein sequence: (SEQ ID NO: 202)
MSKSTINEAVLTQVLNHLRNGQLRRCAEMGLRPEILAQLQQPAVMSILTNTPVSWVDVRVNIDVMEKILATAERSAQEDL
QIERALKLGATTTMIQSFFGLSPEDTATKRLMLEIHPRRGRWRQLDEQIERQIWFRWEHLMQENQVRLEDSMELLDIAMI
LTEEINAGIEQDSPEFISLAIVWSLIQSWLKDGLYPSGKSSQSQAGLQKSQSTLYLASVSSHLPHSAPSATTQVNAETER
QQLLNLVQSEGDTAP.

FIGURE 30R

RL102
DNA sequence: (SEQ ID NO: 81)

ATGAGTATGGCCAAGATCAACCCGCAAGATCTGAAAGACCGGCTACTTGCCCCTGGTTTTACCGCACCGCCCAAGGTTCT
GGAGCAGTTGTCGGACCCGATCAGCGACACGCCCATGAGGCTGACACTCCACGACGTCCTCCCCTGGCACGATAACCCCC
GGACCACTCGAAACCCGAAATACGATGAGCTGAAAGAATCGATCCGACATCGAGGCCTCGATACGCCGCCACCAGTGACT
CGACGCCCTGGAGAGGACAAATACCGCATTCGCAACGGCGGAAACACTCGCCTGGAAATTCTCAACGACCTCTACAAAGA
GACCGGAGACGAGCGCTATTTCAGCTTCGACTGCCTGTTCAAGCCCTGGGACAAGCAGCGCGGCGAAATCATCGCGCTGA
CCGGTCATTTGGCCGAGAACGATCTGAAGGGCGACCTCAAGTTCATCGAGCGCGCGGTTGGGGTGCAGAAGGCGAAATTT
CTTTACGAACAAGAGAACGGCGGTGAAAGCATTTCCCAGCGCGAGTTGGCACGTCGGCTAAAAGCGGACGGCTACCCTGT
ATCTCAATCCCATATCAGTAAGATGTTAGACACTATTGAGGTATTGGCGCCGGCGATTCCTGTGATGCTGTATTCAGGGC
TCGGTAAACCGCAAATCGAGAAACTCCTGTCACTCAGAAAGTCGGCATCCTCCTGCTGGGCACGTCTATACGCTGGTGAA
GGGGTTGACTTCGAAATGCTGTTCCAGGACACCCTGGCAATCTTCGACAGTAGCCCTGACGAATTCATTTTCGAGCGTTT
CCAGGACGAACTCATCGACCAAATGAAGCGCCCCCTGGGCCTGCGTTATGACCAAATCCTGCTCGAGATTACCAACGGGC
AGCAGGAGCAACGCCGCGGCACTCTGGTCGACCTGCCCACACCTGCCGCACCACCTCAACTCCCACCAATTGGGCAGGAA
AACCCTGCTGCGTCGTCTACTGGACAAGCACAAACACAGAGCCCCGCCGCAGATCCCCAAACGTCCAGGACAAGGAGCAA
CCCGGGTAATCCCCTCCCCCCGCCGGCTCCGCCACCACCTGTCCAACAAAAGCAATTGCCCGATGAGGAGCGTGCGGCGG
TCTTGGCAGGCCATATCGTGAGCCCGGTATCGACTAAGATCCAGCAGACTCGCCAACGGCTGGCCGGCCTCGAGGGGGAA
CATCTACCTGTCTTCGATGAAACAGCTCTGCAGGCAATCCCAGTGCAAGTCGGTGGCCTGCACCCGATCACCGATCTCTG
GTACATCGAGCGGTCGATCGATACCCCCGAGATCCTGCGACAGCACATCGCTGATCTGGCTGAAGAGATCGCTCTGCATG
TCGGCGCCCCAGGCGAGATCGTCAGGATTCAGGGCGGTGTGGGTTACACGTATCGCGAGCCCAATGAAGACCATGAGATT
ACTGATTCAGCGCTGCACCTCATGACGCTGCTTCAAGCGGTCAGCGGGCAGGTCCAAGTCGTTCTGAACACTCACGATCA
ACAGACCTGCCGCGATGCACTGGGTGAATTCCAGTTCTCAGCTGGCCTCGCTCAGTTGCTGCTGGGCCAACCCACCACAA
GTGACAAGCCATCCTGCCAGGCAGGCCGTCTCAATGACGAAGCCCTGGTGAAACTGTTCCGGATCATTCGTCTTGCCCGA
CGCCTGGTTGACCTTGAGCTGCCGCCGGCCGCCTCCGAGCAAGCAGCTACTGACCAGTGA

Protein sequence: (SEQ ID NO: 203)
MSMAKINPQDLKDRLLAPGFTAPPKVLEQLSDPISDTPMRLTLHDVLPWHDNPRTTRNPKYDELKESIRHRGLDTPPPVT
RRPGEDKYRIRNGGNTRLEILNDLYKETGDERYFSFDCLFKPWDKQRGEIIALTGHLAENDLKGDLKFIERAVGVQKAKF
LYEQENGGESISQRELARRLKADGYPVSQSHISKMLDTIEVLAPAIPVMLYSGLGKPQIEKLLSLRKSASSCWARLYAGE
GVDFEMLFQDTLAIFDSSPDEFIFERFQDELIDQMKRPLGLRYDQILLEITNGQQEQRRGTLVDLPTPAAPPQLPPIGQE
NPAASSTGQAQTQSPAADPQTSRTRSNPGNPLPPPAPPPPVQQKQLPDEERAAVLAGHIVSPVSTKIQQTRQRLAGLEGE
HLPVFDETALQAIPVQVGGLHPITDLWYIERSIDTPEILRQHIADLAEEIALHVGAPGEIVRIQGGVGYTYREPNEDHEI
TDSALHLMTLLQAVSGQVQVVLNTHDQQTCRDALGEFQFSAGLAQLLLGQPTTSDKPSCQAGRLNDEALVKLFRIIRLAR
RLVDLELPPAASEQAATDQ.

RL103
DNA sequence: (SEQ ID NO: 82)

ATGACCATGGCCCGAGAAACCGAAGATAAGTTCGTTGTCCGTATGCCCTTGGGCTTGCGCGATCAGCTAAAGCAAAAAGC
CGCGGATAACCACCGTTCGGCCAACAGCGAGATCGTCTACCGACTGGAGCGCAGCAACGCGCTCGAAGAAGAACTCGCGC
GAGCAAACCGAATGGTCGACGAACTCTTCGCCAAGAACCAGCGCCTGCAGGCTGAGCTGGCGGCGGCGAACACGCCTCAG
GTGGCGGAGGCATGA

Protein sequence: (SEQ ID NO: 204)
MTMARETEDKFVVRMPLGLRDQLKQKAADNHRSANSEIVYRLERSNALEEELARANRMVDELFAKNQRLQAELAAANTPQ
VAEA

RL104
DNA sequence: (SEQ ID NO: 83)

ATGCCTATCAAACACGCCATCGTCCACCTGATCGAGAAGAAGCCCGACGGCACCCCGGCCGTGCTCCACGCGCGCGACGC
CGAGCTGGGCGACTCCCAGGCCATCGAGAACCTGCTGGCCGACCTCAACGAAAGCTACAACGCCAAGAACAAGGCCTGGG
GCTTCTTCCAGGGCGAGTCCGGGGCCTACCCGTTCAGCGGCTGGCTCGGCGAGTACCTGGAGGGCGACCGCGACTTCGTC
GGCTTCAGCCGCGAAGCGGTCGAGCACCTGCAAAAGCTGATGGAGGAGTCCAATCTCTTCACCGGCGGCCACGTCCTGTT
CGCCCACTACCAGCAAGGCATGACCGACTACCTGGCGATCGCCCTGCTGCACCACAGCGAAGGCGTGGCGGTGAACGAGT
CGCTGGAGGTCACCCCGTCGCGCCACCTGGACCTCGGCCAGTTGCACCTGGCCGCGCGGATCAACATTTCCGAATGGCGC
AACAACAAGCAGTCGAAGCAGTACATCTCGTTCATCAAGGGCAAGGGCGGGAGGAAGGTCTCCGACTATTTCCGCGACTT
CATCGGCTGCCAGGAAGGGGTGGATTCGCCGAGCGAGACGCGCACCCTGCTGAAAGCCTTCAGCGATTTCGTGGAAAGCG
AGGACATGGCCGAGGAACAGGCCCGCGAGAAGACCGAGACGCTGGTCGACTACGCCACCTCGCAGGCGCGCATCGGCGAG
CCGATGACCCTCGACGCGCTTTCGGAACTGATGGACGACCAGCAACCGCGGGCGTTCTACGACTACATCCGTAACAAGGA
CTACGGCCTGTCGCCGGAAATCCCGGCGGACAAGCGCACCCTCAACCAGTTCCGCCGCTTCACCGGCCGCGCCGAAGGCC
TGTCGATCAGCTTCGAGGCGCACCTGCTGGGCTCCAGGATCGAGTACGACGAGGAGCGCGACACGCTGCAGATCAGCAGC
CTCCCCACTCAACTCCGCGACCAGCTCAAGCGGCGCAAGGCCCAAATTGGAGAATGA

FIGURE 30S

Protein sequence: (SEQ ID NO: 205)
MPIKHAIVHLIEKKPDGTPAVLHARDAELGDSQAIENLLADLNESYNAKNKAWGFFQGESGAYPFSGWLGEYLEGDRDFV
GFSREAVEHLQKLMEESNLFTGGHVLFAHYQQGMTDYLAIALLHHSEGVAVNESLEVTPSRHLDLGQLHLAARINISEWR
NNKQSKQYISFIKGKGGRKVSDYFRDFIGCQEGVDSPSETRTLLKAFSDFVESEDMAEEQAREKTETLVDYATSQARIGE
PMTLDALSELMDDQQPRAFYDYIRNKDYGLSPEIPADKRTLNQFRRFTGRAEGLSISFEAHLLGSRIEYDEERDTLQISS
LPTQLRDQLKRRKAQIGE.

RL105
DNA sequence: (SEQ ID NO: 84)

ATGCGTAGTTTCCTTCGCGGCGCCCGGGAAAGCGTTCGCCGGCTGGTGGCCTTCGCTCAAGCAGAAGGCTGGAGCGTCGA
CCGCTCCGCAGGCGGCCACTTGAAGCTCAGCAAGATCGGCTGCGCCTCGATCTTCATTTCTTCCACGCCAAGCGACGCAC
GCGGCGAGCTCAATGCCCGCGCCCTGCTCCGTCGAGCCGACAGGCAGCGTTCCCTGAACCAGGAGTCTTTCTGA

Protein sequence: (SEQ ID NO: 206)
MRSFLRGARESVRRLVAFAQAEGWSVDRSAGGHLKLSKIGCASIFISSTPSDARGELNARALLRRADRQRSLNQESF

RL106
DNA Sequence: (SEQ ID NO: 85)

ATGCCTGACGTCACAGCCTACCGGCCGCTCGAGCACTTCCAGAAAGTCGAGCTGATGCTTGAGCTCAAGTTGCGTGAAGG
TCCTTCGTGGATCTGTCTGAACTGCGGCTATCACCTGGATGGCAGCGGCGCACAGCCCTGCCCTGACTGCGGAAAGTCGC
GCTACTGGACCAGCGGTTGGAGTGTAGGTCGTGGCCATCGCTTCTCGGCAGCAAGGGAAGAGTGGGAAAACCGCCTCAGG
ACACGGTCGCGGTCACCTGTCGCGTCAACGGCACCAGTAGCAACTGACGACGTATGCACTCAACTGCGCACAGAGGTCCG
CATGCTGCGTTCCGCGCATGACGACCTGGCCTGCAGCCGGCAGAGCGATCGTCGCAGCCTTCAGGCGCTGGTGAAACGTC
TCCTGGATGCCGCCGCCACCGATAGCCTTCCCCGCTCCCTTGCAGAGATGGAGACCTGGCTGCAGCTCAACAGCGAGGAG
ACCACGAATGCGTAG

Protein sequence: (SEQ ID NO: 207)
MPDVTAYRPLEHFQKVELMLELKLREGPSWICLNCGYHLDGSGAQPCPDCGKSRYWTSGWSVGRGHRFSAAREEWENRLR
TRSRSPVASTAPVATDDVCTQLRTEVRMLRSAHDDLACSRQSDRRSLQALVKRLLDAAATDSLPRSLAEMETWLQLNSEE
TTNA.

RL107
DNA sequence: (SEQ ID NO: 86)

ATGAAGGCGTCCCAGACCTATCAGTGCATCGTCAAGTTCGATGGCGCCGGTTTCTGGACCAATACCATTCAGAAGCAGCG
TGCGACCTGCACCTGGAGCGACAAGGTGGCAGCCTCCCGCCTTGCCGAACGACTGTTTGGCGAGGACAACGCATACATCA
CCCGTATGCCGGTACAGGCAGGCGACCACGAAAAGCGCATCGAGAGCCGCTGGGCGCTGTCCTGTAGAAATCCCAAGGAG
GTAGCGCGCGATGCCTGA

Protein sequence (SEQ ID NO: 208)
MKASQTYQCIVKFDGAGFWTNTIQKQRATCTWSDKVAASRLAERLFGEDNAYITRMPVQAGDHEKRIESRWALSCRNPKE
VARDA

RL108
DNA sequence: (SEQ ID NO: 87)

ATGAACACTGAAGCCCGCTTTCCGAGTATCCACGCCTCGGCCGCGTTCACCGACTCGGCAGTGGTTCATGCCAATCACGT
TGGGGTCAACCCCATCGAGCTGGACGCCCTCAGCCAAGTGATCTCGCGCCTTTCGCGGGACGAGAGCACGGTCGCACCCA
GTTCGATGGAGCGAGAGCTTCGTGAGCTGGAGGAACTGGGGTACATCGAAATCTCGACCACCCAGGCCGGGACTCTGGTG
GTCACTACGCGCGCTCCGGGGCAATTGCTTTCGGCTTACTTCTGGTCGGTATGGATCCCGCGACACCTGTTCAGCTGCTC
GCTGAAAGTGAGCCTGGTGCCGCACCTCTGCTGCGGCACTCAGGACTCCCAGCACCTCACCGCCGTGTTCCGCATTGCAG
GCAGCAAGGACGCCGCGCGCGAGTTCCTGCATCAGTTGGCCAACAACTATCCCGGGCATGAGCCGGAGTTGCCCGAACTG
GTGGCCGTTCAGGTCGGTGATGCACTCAGCAAGGAGGCCGAGTCATGA

Protein sequence: (SEQ ID NO: 209)
MNTEARFPSIHASAAFTDSAVVHANHVGVNPIELDALSQVISRLSRDESTVAPSSMERELRELEELGYIEISTTQAGTLV
VTTRAPGQLLSAYFWSVWIPRHLFSCSLKVSLVPHLCCGTQDSQHLTAVFRIAGSKDAAREFLHQLANNYPGHEPELPEL
VAVQVGDALSKEAES.

FIGURE 30T

RL109
DNA sequence: (SEQ ID NO: 88)

GTTTCTGCAGATCAGCCGGGAGGACTCGAAGTGAGTGTTCTGGAACTTACGCCGCCGCACTCCGTCGAGGCGGAGCAAGG
GGTACTCGGCGGCCTGATGCTGGACAACGCGGCATGGGACATTGTCGGCGATCAGTTGCAGAAGGAGGACTTCTTCCGGC
ATGAGCATCGGCTGATCTTCACCGCCATCAGCGAGTTGGCCGCGAAGGATGCTCCGTTTGATGTCGTGACTGTGTCGGAA
GCGATCGAAGACCTTCCAGAAGCTGGCGGGCTGGCCTACCTCGGCCAGCTCGCCGACAACACGCCCTCCGTGGCCAATAT
CGAGGCTTACGCGCAGATCGTTCGCGATCGGGCACACCTGCGGCAGCTGATGTCTCTCGGGCACCACTGCACCAGGACCG
CCTCGAACCACCAGGCAAATCCCTCTGAGGTTCAGGAGGAGATTGAGCAGAAGCTGTTCGGCCCTTGGCCAGGACCACCA
CAACGCCGATTTCGTCGATATCAACAAGAGTCTCACGAAGATCGTCGACACCATCGATTACCGCTTCAACAACAACGTGA
CGGTAACGGGGGTCCCGACTGGCCTGAAGGATCTCGACGCACTCACCGGCGGACTACAGAAGTCGGATCTCATCATCGTC
GGTGCCCGCCCCGCGATGGGCAAAACGTCGTTTGCCCTCAACCTGGTCGACACCGCGCTCCAGAGCGACCAACAGAAGTC
TGTTCAGGTGTACAGCATGGAGATGCCGGCAGAGCAGTTGCTGTTCAGGCTTGCCGCCCTGTTCGGCCACCTGGACCTGG
GCAAGCTGATGAAGGGCCAACTGCAAGAAGAGGATTGGCCCAGACTGTCTGGCGCGATCCAGCGCATAAACGACTATGGC
AGCCGGCTGGTCATCAACGATCAGGGCAACCTCACGCCGACAGAGCTGCGCGCCAAGGTTCGCCGGGCGGCCAGGAAGTA
CGGACACCCCGCGCTGATATTGGTCGACTACCTGCAACTGATGAGTTGCCCAGGCCTGGAGAATCGAGCCACCGAGATCT
CGGAAATCTCCCGCTCGCTGAAAGCGCTGGCCAAGGAGATGGACTGTCCCGTCGTAGCTCTATCCCAGCTAAATCGCGGC
CTAGAGAACCGGACGAACAAGCGACCGAACTGCGCGGACCTACGAGAGAGCGGCGCAATCGAGCAGGACGCGGACGTGAT
CATGTTCGTGTACCGCGACGAGGTCTACCACCCCAACACCGAGGCCAAGGGCATCGCCGAAATCATCATCGGCAAGTATC
GCAACGGTCCGATCGGCACCGTCCACACCGCCTTCATCGCCAACCAGACCCGCTTTGCCGACCTGGCGCCGGGGACCTGG
CAA

RL110
DNA sequence: (SEQ ID NO: 89)

ATGACTCGCTCTGCTCTCTCGACCATCGCCTACGAGGCCCTGGTGCGTGCCCGCCGCAAATTCAGCAACCGAGAGGAGCG
CTGCATCCGCGAAACCTGGACCGCCGAACAGGAACTGGTGCTGCTGCGCCTGTATCCGGATATGCCGAACGAGGTCCTGG
CAGCCAGGTTGAACAAAACGCTCCAGCAGATCTGCTCCAGAGCGTATCGGCTCGGGCTGAAAAAAAGCCCTGAGTTCTCC
AAGAAGATCCGGCAGGACTGGGGCAGCGCAACTCGGTTCAAGAAGGGAAACACCCCATGGAACTGCGGCATGAAGGGGCT
GCCCGCGCGAGGACGCGCACCAGAAACGCAGTTCAAGAAGGGGCAAAAGCCCCACACATGGCTCCCAGTCGGCAGCACGC
GGGTCAGCGCTGATGGCTACCTGCAACGAAAGATCTCGGATACCGGCTATCCCCCCCGGGACTGGAAGGGCATCCACATC
CTGCTCTGGGAAGAACACTTCGGCCCCATCCCAACCGGCCATTGCGTCTGCTTCAAGGACAACAACAAGCAGAACGTCGT
CATCGACAACCTGGAGCTCATCACCCGGGCCGAACGCATGCGCCGCAACTCCATCCATCGCTATCCACCTGAGCTGAAGA
GCGCAATCCGCGTCATCAGCAAGCTCAAACGCACCATTCAGGAGGTCGAGCATGAAGAACAAGATTGA

Protein sequence: (SEQ ID NO: 210)
MTRSALSTIAYEALVRARRKFSNREERCIRETWTAEQELVLLRLYPDMPNEVLAARLNKTLQQICSRAYRLGLKKSPEFS
KKIRQDWGSATRFKKGNTPWNCGMKGLPARGRAPETQFKKGQKPHTWLPVGSTRVSADGYLQRKISDTGYPPRDWKGIHI
LLWEEHFGPIPTGHCVCFKDNNKQNVVIDNLELITRAERMRRNSIHRYPPELKSAIRVISKLKRTIQEVEHEEQD

RL111
DNA sequence: (SEQ ID NO: 90)

ATGGACAAGCAAAAAGTCCTCGCCAAGGTCGAGAAGCTGATGGCCCTGGCGAATGCCAAGGGGGCCACGCCGAACGAGGC
GGAAACCGCATTGCGCCAGGCCGCGATCCTGAAGCGGCAGTTCGACCTCAGCGATGCGGAGATCTCGGCCCACACGGTGG
AAACCGCGTGCGTTCCCACTCGAACCAGGCGCTCTCCTGCCCCATGGCTGCATGAACTGGCCGGGATCTGCGCCAGTTCC
TTCGGCTGCGACTACCTGGCGGCATACGCGATGCCAGCGGGCTGGACGTTCAAGTTCATGGGCCGAGGGATCGGCCCTGA
GCTGGCCGCTCACGCCTACTCTACGCTCCACCACCAACTGGTGGCAGCGCGCTCGGCTCATGTCGCCCAACAGAAGCGCT
GCAAGCTGTCGACCAAGCGTCGTCGCAGCAAGCTCTTCGTCGAAGGCTGGCTTCTCGCAGTGCGTTCGCTGGTACGTGAA
TTTGCTGGCAGGCCGGACGAGTCGACTCAAGCAGCCATCAAGGCCTACCTCGAACTACACCATCCGGCGTTGAAGTACCT
GGAGCCGGCGGCGCTTACGAAGGCCCTTGCCTATGACCAGGCCTCGCTGCAAGCAGGCTGGGAGCACGGCAAAAACACTC
GCCTGCACCGCGGTGTCAGCCGGCGAGTTCAGGGCGCGCTCGAGCAGGGAGGTTCCCAATGA

Protein sequence: (SEQ ID NO: 211)
MDKQKVLAKVEKLMALANAKGATPNEAETALRQAAILKRQFDLSDAEISAHTVETACVPTRTRRSPAPWLHELAGICASS
FGCDYLAAYAMPAGWTFKFMGRGIGPELAAHAYSTLHHQLVAARSAHVAQQKRCKLSTKRRRSKLFVEGWLLAVRSLVRE
FAGRPDESTQAAIKAYLELHHPALKYLEPAALTKALAYDQASLQAGWEHGKNTRLHRGVSRRVQGALEQGGSQ

FIGURE 30U

RL112
DNA sequence: (SEQ ID NO: 91)

ATGAGTGACCCCAAGCTCAAGCCCTGCCCGCTCTGCGGCAGCACGAACATTCGAATGCTGGAACCCGAGCTGCTCGACAC
CGATGCCTGGAACTGTGCCATTGAATGCCTGGACTGCCAGGTTCACATCGGGCCGTCCTACTGCGAGCCAGACCCGGTAA
CAGCGAGGTATTCAGCACAGATCGACTGGAATAGACGCCCAAGCGCAAAAAACCACGCGGACGAGCGTGAGCAGTTCTTG
ATGGCCAACCTGCTCGCCGCCCTGGAGGTCGCACTGGGCGACGTAGCAGCCCTGGCTATTGTCGATCGGGTAAGACAGGC
CACAGACCGAATTTACCCAACTTCGAACCTCTCCCCTGTTCCGCAGGCCTGGCTCGATGTACAGGCCGAGCGCCGGCGCC
AGATCACCGTCGAAGGTTTCGATACCAGCAACGACGACGCTAGCGCTGGCCTGATCGCCCTGGCGGCCGGCTGCTACGCG
CTCCATGCCGGCGGCATCGGCACCGACTGGCCGGGCGGCATTCGGAATGGCTCTGCACTGTTCTGGCCCTGGGACGAAGA
GTGGTGGAAGCCTAAGTCGGCGCGCGAGAACCTGGTACGCGCCGGCGCCCTAGTGCTGGCCGAGATCGAGCGCCTGGACC
GCTCCGCCACCGAGCAGGGCTCAACCATCTGCAAGGGGGGCGCGTAA

Protein sequence: (SEQ ID NO: 212)
MSDPKLKPCPLCGSTNIRMLEPELLDTDAWNCAIECLDCQVHIGPSYCEPDPVTARYSAQIDWNRRPSAKNHADEREQFL
MANLLAALEVALGDVAALAIVDRVRQATDRIYPTSNLSPVPQAWLDVQAERRRQITVEGFDTSNDDASAGLIALAAGCYA
LHAGGIGTDWPGGIRNGSALFWPWDEEWWKPKSARENLVRAGALVLAEIERLDRSATEQGSTICKGGA

RL113
DNA sequence: (SEQ ID NO: 92)

ATGAACCTCCAGAACCGCAACAACCTCCTACTGAGCTTGATCGCCGAGACCCAGTTCGACGCCTACGTGCAAGGCTACAT
GGCCAAAGCAGGCGCTGCCGCCGGTGCTTCCGAGAATCTGCAAATCGAGGCTGAAGGTGCTGCGATGTTGCAGGGCCTGG
TCGCTCCGGTTCGCGCTCAGCAGCGTGCCTGTGGACAGTCCCTGCAGAACGCACTGCTCCAAATCGCCCACGACCTACTG
TTGCAGACGAAATCGCAACTGGCCATCGCGGCCAATGCCAGTTCGATCCAAGTGATCCAGCGGGACATGAACAGGGCGAT
CTGGAACATAGCTACTGCCATCGATCACCTGGCCGAGTTCGCCCAACCCTCGCAGGACACTGTGAGGGTCATCGAACGGC
TGATGCTCTTCGTCGGCAGCTCATCAAGCACTGAAGGCCAGCAACTGGCCGCCGAGGCAAATGCGGTGCTCGGCATGAGC
GTGGGAGGCCTGGCATGA

Protein sequence: (SEQ ID NO: 213)
MNLQNRNNLLLSLIAETQFDAYVQGYMAKAGAAAGASENLQIEAEGAAMLQGLVAPVRAQQRACGQSLQNALLQIAHDLL
LQTKSQLAIAANASSIQVIQRDMNRAIWNIATAIDHLAEFAQPSQDTVRVIERLMLFVGSSSSTEGQQLAAEANAVLGMS
VGGLA

RL114
DNA sequence: (SEQ ID NO: 93)

CTGAACAAGTTCGGCAGCGCCGCCGACCTTCGGAGCCAGCAGGCCAAATTGACCGGCGCTACGCGAGAAATACGCAAGCT
GACTGGTGGCGGTATCGACCTGTTCGGGAAGCTGGGTTGCTACTTGAGCTTCGAACAAAAGCAGCTCCTACAAGACGCAG
CGCGCTTGCTCGACTCGGTGAACAAGCAGATCGAGCATGCGAAGGAAAAGCGTGATCGCTACGAGAAAAAAGCCAAGAAG
CGGCGCGAGCTACGTGAGCGCCTGGCCAAGCAACTGGTCGCCTCGAACTACCCGCTTCCGGGAAATACGCTCGAAGATCG
GCTGGAAATCCTGCAGATCGCGTTGATCTACAACCGGGCCAGGGTGTTCGATCACCTGTACTCCACGCACCAGCTCCACT
CAAAACTCAAACGCTGGCTGGAGCGTCCAAAGCAGCTCATCGGATGGCGCAGTGAAGCCGAGTATTTCGCTAGTCAGGTG
GGGAGCCTGCGATGTGACTTCATTAGCCATCTGACTAACGAAATCGCGTACGACGATGGCAGTGAAGTCGAGGAGCGCCT
GCGCGTCATCAAGCAGAAGGTCGCTGACTGCACCGCACAGATCGCTCTGACCAGCGAGGAGCAGGAAACCCTTCGGCTCT
GGACAGACGCTCTGCAATCGGCTCCGGAGGGCCTCATATGA

Protein sequence: (SEQ ID NO: 214)
LNKFGSAADLRSQQAKLTGATREIRKLTGGGIDLFGKLGCYLSFEQKQLLQDAARLLDSVNKQIEHAKEKRDRYEKKAKK
RRELRERLAKQLVASNYPLPGNTLEDRLEILQIALIYNRARVFDHLYSTHQLHSKLKRWLERPKQLIGWRSEAEYFASQV
GSLRCDFISHLTNEIAYDDGSEVEERLRVIKQKVADCTAQIALTSEEQETLRLWTDALQSAPEGLI

RL115
DNA sequence: (SEQ ID NO: 94)

ATGAATGCGAAAGCGACTTCGGTTGTATCCACCAAGGGTGGTGTAGGAAAATCCACCACCGCCGCCAACCTCGGTGCATT
TTGCGCCGATGCAGGCATACGAACCCTCCTCATCGATCTGGACCCCGTCCAGCCCTCCCTATCCTCGTACTACGAGCTGC
CGGAAGTTGCCCAGGGCGGCATTTACGACCTGCTCGCCGCCAACATAACGGACCCGGCGAGGATCATCTCCAGGACGATT
ATCCCCAATCTGGACGTCGTGATTTCCAACGACCAGAACAATCAGCTCAACAACCTACTGCTCCAGGCGCCCGATGGCCG
GCTACGCCTGGCGAACCTGATGCCCGCTCTGAAAGAAGGCTACGACCTGGTGCTGATCGACACCCAGGGTGCGCGCTCAG
CTTTGCTCGAAATGGTTGTGCTTGCATCGGACCTGGTTGTTTCCCCCCTCCAACCCAACATGCTTACCGCCCGTGAGTTC
AACCGCGGCACCATGCAAATGCTCGACGGCCTACGCCCCTATGAGCGTCTCGGCATGCGGATCCCCAATGTTCAGATCGT
CATCAACTGCCTGGACCAGACCAATGACTCCCGGGCAATTCACGAGAATGTGCGTGCCATCTTCGATGAGCATCAGGACA
TTTCTGTGCTCGAAACGACTGTCCCGGATGCCGTCGTGTTTCGCAACGCAGCATCGCGCGGGCTACCAGCGCACCGCCTC
GAAACGCGGCAACCCTCCAATCGCACATCAGCGCCCGCGCTGGAAATCATTCGAAACCTGGCCATCGAGGTCTTTCCCGA
GTGGACTGACCGCTTCCTGGCGCTGACGCCGGGAGGCGGTTGCAGCACTGGTCAAGGGAGGGCGCTGACATGGCGAAGAC
TCCTATCACCCAAGCCCGCGACGTCGACGCGGAACTTGTGCTGGAACTGA

FIGURE 30V

Protein sequence: (SEQ ID NO: 215)

MNAKATSVVSTKGGVGKSTTAANLGAFCADAGIRTLLIDLDPVQPSLSSYYELPEVAQGGIYDLLAANITDPARIISRTI
IPNLDVVISNDQNNQLNNLLLQAPDGRLRLANLMPALKEGYDLVLIDTQGARSALLEMVVLASDLVVSPLQPNMLTAREF
NRGTMQMLDGLRPYERLGMRIPNVQIVINCLDQTNDSRAIHENVRAIFDEHQDISVLETTVPDAVVFRNAASRGLPAHRL
ETRQPSNRTSAPALEIIRNLAIEVFPEWTDRFLALTPGGGCSTGQGRALTWRRLLSPKPATSTRNLCWN

FIGURE 31A

RS01
DNA sequence: (SEQ ID NO: 95)

ATGGGGATCTACCGCCCGAAGCAGTCTCGCCTAGCGATACCGATACTGAGGGGCCGGCTACCGGACGAAAGGTAGCTGCG
CCTCCCAGCAGTTCGCTAGGCCTGTAGGAAAAATCTGGAATTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCA
GAGCCCCGCAATTTCAAGGCCGAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGAGGGCACATTGCTACGCCTGCA
GAATGGTTTCAGGGCCTGAAAAACAGAAAAGCCCACCTAAATAGGCGGGCTATTCCATATTGACATCACGTCAATGCGGG
CC

RS02
DNA sequence: (SEQ ID NO: 96)

ATGACGCCGCAGCAGCTCACCGAGGAGTACATCTTCGCGCACGATCTCCGAGAAGCCAGCGCGAAGATCTACCGCGCCGC
GACCAAGGCGCTGCTCAAGCACTTCGGCCCTACGGCAACCGTACAGGAGGTGGACCACAGGTCTGTGCTGGGATGGCGGC
GCAAGGTCCTGGAACAAGGCCTGTCGAAGCGGAGCTGGAACACGTATTCGAATCATCTGCGAACGATCTGGGGCTATGCC
ATCGAGCACGAGCTGGTGACACACTCCCAAGTCAACCCGTTCAGAAAGACCACCGTCATCCCCCCCAGGCGAGCAAGCAA
AACCGTCGCAGCCGAAGCCATCCTGCGCGCCCGCAATTGGCTCAACATGCAGGTCGGCGCCGAGCGCTGCACTGGCGATC
GCGCACGCATCACTCCCGCCTGGTTCTGGCTTTGCACGTTTGAGGTCTTCTACTTCACCGGCATCCGGTTGAATGCGCTG
TTGTGCATCCGCAAGCGCGACATCGACTGGGAAAATCAACTGATCCTCATCCGCGGCGAGACAGAGAAAACTCATAAAGA
GTTCGTAGTGCCAATAACGGAGGGGCTTGTGCCTCACCTATCGCGGCTCCTGCAGGAGGCCGATAGAGCCGGATTCGCCG
ATGACGACCAGTTGTTCAACGTCAACCGGTTCTCACCGCACTACAAGAGCAAGGTGATGAACTCCGACCAGGTCGAAGCC
ATGTACCGGAAGTTGACCGAGAAGGTTGGGGTGCGGATGACTCCGCACCGTTTCCGGCACACCCTGGCCACCGACTTGAT
GAAGGCACCCGAGCGGAACATCCACCTCACGAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCG
AGGCCGACTACGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGTCAGGAAGGTG
GATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCAACCTCTCGCTCGAATGGGTGAAGCGCCGCC
ACAGGAGGCTAGGACAGAACCTGCAGAACCAAGGGAGCACACACCAGGGACAGGCATTCAGGGAGATGCAACCGCGTGGG
AAGAAGCGCTACCACAGCCACCTGACACCTTCGAGCAAAGCGTGCTGTTCACTCTGATGGCTCAACACCTATCGAACCGT
GCCGCCACGGCCTCCGCGGCTTCCACCGCAACAAGCGGATCTGGAGGATGGGGATCTACCGCCCGAAGCAGTCTCGCCTA
G

Protein sequence: (SEQ ID NO: 216)

MTPQQLTEEYIFAHDLREASAKIYRAATKALLKHFGPTATVQEVDHRSVLGWRRKVLEQGLSKRSWNTYSNHLRTIWGYA
IEHELVTHSQVNPFRKTTVIPPRRASKTVAAEAILRARNWLNMQVGAERCTGDRARITPAWFWLCTFEVFYFTGIRLNAL
LCIRKRDIDWENQLILIRGETEKTHKEFVVPITEGLVPHLSRLLQEADRAGFADDDQLFNVNRFSPHYKSKVMNSDQVEA
MYRKLTEKVGVRMTPHRFRHTLATDLMKAPERNIHLTKCLLNHSNIQTTMSYIEADYDHMRAVLHARSLAQGALENVRKV
DYSGSPQASAKPKPCGQPLARMGEAPPQEARTEPAEPREHTPGTGIQGDATAWEEALPQPPDTFEQSVLFTLMAQHLSNR
AATASAASTATSGSGGWGSTARSSLA.

RS03
DNA sequence: (SEQ ID NO: 97)

ATGAAATCTGGTATCGCGACCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATACCGTCGACGGGACCGCCAT
GCTGGTCACGCCAGGAATCTTCAAGCGTTATGTCCAGGAGCATCCGGAGGTTGAAAAGCTGGCCCAGGCCAAGGAGACCG
CCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAGAAACAGGGTCTTCACCGAAAGACCAGTAAGAACCTGAATATCTGGACC
ATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGTGCA
GCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA

Protein sequence: (SEQ ID NO: 217)

MKSGIATRRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPEVEKLAQAKETAGWKLVQRAFEKQGLHRKTSKNLNIWT
IKVSGPRKTKELKAYLLQDPKLLFPVQPLDNPSLTVITDAEGGVE

FIGURE 31B

RS04

DNA sequence: (SEQ ID NO: 98)

ATAGACCAGTTGAGTGAGCAGGAGTCGGTGGAAGTGGTCTGCTCAGCTTTCGATGTGGCGCGGTCTTGCTACTACGTCCA
CCGTCTTCGACGGCGGCGTGTCGATGCTCGCCGCGTGGCGCTACGCAGCCAAGTCAACCAGTTGTTCAGCCAGAGTCGGG
GCTCGGCCGGCAGCCGCAGCATTCTGGGCATGCTGCGCGAAGAGGGCGTGACCATCGGCCGTTTCCGAGTGCGTCGGTTG
ATGCGTGAGCTGGGCCTGGTCAGCAAGCAACCGGGCTCGCACGCCTACAAACAGGCCACGGTTGAGCGGCCGGATATCCC
GAATCGGCTGAACCGCGAATTCGCGACCGAGCATCCCATACAGGTGTGGTGTGGCGACATCACCTACGTCTGGGCGCAAG
GCCGTTGGCACTACCTGGCCGCGGTGCTGGATCTGCTGATCGGCTGGGCGTTCTCGGCCAAGCCGGATGCCGAACTGGTG
ATCAAGGCCCTGGACATGGCCTACGAACAGCGCGGCAGGCCACAGCAGGTGCTGTTCCATTCAGACCAGGGCAGCCAGTA
CGCCAGCCGCCTGTTTCGGCAACGGCTCTGGCGCTATCGGATGCAGCAGAGCATGAGCCGTCGGGGGAATTGCTGGGATA
ACTCGCCGATGGAGCGCCTGTTCCGCAGTCTGAAGTCGGAGTGGGTCCCGTCAACGGGTTACCTGACGGCGCAGGAGGCC
CAACGGGACATCAGTCATTACTTGATGCACCGCTACAACTGGATCAGGCCGCATCAATTCAACGACGGGTTACCACCTGC
GGTGGCCGAAGAAAAACTCAACCCACTGTCCGGGATGGGTTGA

Protein sequence: (SEQ ID NO: 218)

IDQLSEQESVEVVCSAFDVARSCYYVHRLRRRRVDARRVALRSQVNQLFSQSRGSAGSRSILGMLREEGVTIGRFRVRRL
MRELGLVSKQPGSHAYKQATVERPDIPNRLNREFATEHPIQVWCGDITYVWAQGRWHYLAAVLDLLIGWAFSAKPDAELV
IKALDMAYEQRGRPQQVLFHSDQGSQYASRLFRQRLWRYRMQQSMSRRGNCWDNSPMERLFRSLKSEWVPSTGYLTAQEA
QRDISHYLMHRYNWIRPHQFNDGLPPAVAEEKLNPLSGMG

RS05

DNA sequence: (SEQ ID NO: 282)

ATGAGCAAGCAACGACGTACGTTTTCCGCCGAGTTCAAACGAGAGGCCGCGGCCCTGGTGTTGGACCAAGGCTACAGCCA
TATCGACGCCTGCCGTTCGCTGGGGGTGGTGGATTCGGCCTTGCGCCGTTGGGTGAAGCAGCTCGAGGCGGAGCGCCAGG
GTGTGACCCCGAAGAGCAAGGCGTTGACGCCTGAGCAGCAAAAGATCCAGGAGCTGGAAGCCCGGATCAACCGATTGGAG
CGGGAGAAAGCGATATTAAAAAAGGCTACCGCTCTCTTGATGTCGGACGAACTCGATCGTACGCGCTGA

Protein sequence: (SEQ ID NO: 219)

MSKQRRTFSAEFKREAAALVLDQGYSHIDACRSLGVVDSALRRWVKQLEAERQGVTPKSKALTPEQQKIQELEARINRLE
REKAILKKATALLMSDELDRTR

RS06

DNA sequence: (SEQ ID NO: 99)

ATGTTGTATTTTTCTTGCAGTATGAAGATGGGTGGTTGGGTCGGATATAGGTACTTCTCTCTATTTTCTTTAATTGCTCT
CATCTATGGGTGTGTCGGTGGTGGAGGTGGATCGGATGAGATTGGGCAGCACTGCTTTGAGAGAGAGCAAAAGCTTTCCG
GAGTTAATGATAATGAAGAGGGGAGTGTGAGGTTGAATCGGCTGAACTGCGATCCAATTGAAGGTCGTGTTCTTGAATCA
GAGAAGCTGATAAGAAAGCCGCCCAATGAGCTGGGTATTCACTGA

Protein sequence: (SEQ ID NO: 220)

MLYFSCSMKMGGWVGYRYFSLFSLIALIYGCVGGGGGSDEIGQHCFEREQKLSGVNDNEEGSVRLNRLNCDPIEGRVLES
EKLIRKPPNELGIH

RS07

DNA sequence: (SEQ ID NO: 100)

ATGAAAAAATCACTTGTTATGTCGGCTGTGCTTTTGGTGGCTAGCAATTTCGCGTGTGCTGATGAGGGCTCAAATGATGG
AAGTGAGATATGTCGGGCGCAGGGTGGAGTTGAAATAACAAGTCTGGGGGAAGTCTCAAAGGGTGTGGATGTTGAAGATG
TTGTAGTTTGTTCGATTCTTCCAAGTAATATGAAGTCGAGTCAAAGAGCGCCTACACTCCCTCCTCTGCAAAGGATGATC
ATTTCGGCAATGCCTTCACCAGGAACGGTCACTGTTTCTGCCAGCGGAGATAGGAAATTTACAACATCTTGCCGGGCAAA
TCTTTATGCTCCACGTTATGCCAATTTCTATCCAGACGGTGTTAGCAGGGGAACATCAGATCTACGATGTGTTGGTTACA
ATACACCCGGGAATTCATCTCAAGGGTGTAATGTGTCATGGGACGGCCCGACCGACATTCAATTGGGTGTTGAGCCATAT
GGCGGATCTGTTGTTGTAAACTACAGTTGCACTGCATTCAAAACAACGATTCCAGTGATAATGAGCTACAGTTATCGTGA
TGGGCGGGCAGTGTATGGCGAGGTCCAGAATGTGTCAGGAATAATAAATGTGGTTTTGAACTAA

FIGURE 31C

Protein sequence: (SEQ ID NO: 221)

MKKSLVMSAVLLVASNFACADEGSNDGSEICRAQGGVEITSLGEVSKGVDVEDVVVCSILPSNMKSSQRAPTLPPLQRMI
ISAMPSPGTVTVSASGDRKFTTSCRANLYAPRYANFYPDGVSRGTSDLRCVGYNTPGNSSQGCNVSWDGPTDIQLGVEPY
GGSVVVNYSCTAFKTTIPVIMSYSYRDGRAVYGEVQNVSGIINVVLN

RS08

DNA sequence: (SEQ ID NO: 101)

ATGCTTATTAAAATTCTTCGAATTATATTCTTGTTGCCTATAGTTGGTTTGGCACAGCAGGCTGCTGCCTCCCCGCCCGC
AGAGTCACACTCGGAACAATCTGAATCTTCGTGTATCGATGTCCAAGTCAATGGAGCACGTAGCCTGTCTTATAACTGCA
TGGCTCAGCAAATGACTCCACCCAAAGAGGATCCTCGGCGTCGGAACCCTACCTTGAACTCCACATTAGCGTCTGAACGC
GCCACTCGCCTGCCACCCACACAGACAGGACTTTTTACCAGCCTTCATCAACGTGCCATATCGAACTCGAAAGACTAG

Protein sequence: (SEQ ID NO: 222)

MLIKILRIIFLLPIVGLAQQAAASPPAESHSEQSESSCIDVQVNGARSLSYNCMAQQMTPPKEDPRRRNPTLNSTLASER
ATRLPPTQTGLFTSLHQRAISNSKD

RS09

DNA sequence: (SEQ ID NO: 102)

GTGAGTAGTACTAAGAGTAAGCCGATAGCCAGGGGGCGTGGTGGCCCATTTGGGGAAGTGATGAAGAGGTGCGGGCTTGT
ACCGGTTCGAGGAAGGAATAGACAGCAGACAGGATCGCTTGCGATGGGGCAGCAGGAAACCATCAGCCCGTCCGTATCCA
GAACTGCTGCTTGCAGCGTTAGGGGTGACTCCCTCATGCCCTAG

Protein sequence: (SEQ ID NO: 223)

VSSTKSKPIARGRGGPFGEVMKRCGLVPVRGRNRQQTGSLAMGQQETISPSVSRTAACSVRGDSLMP

RS10

DNA sequence: (SEQ ID NO: 103)

ATGGAACGCTTGCTCGAGAGCATTTACATCAATGCCCGGCCGGCGATGGAGTTGAGGCTTAGCCTCACCAGCTCCGGCCG
CAAGAGAATGGTAAAGATTGTGGATGGGGAGGAGGTCGAGGTTCTGCCAGGTGAAGTGCAGGGCATCCTGGAGGCCCAAA
AGAGGGATGTTGGAATCCTCGCCGACTTCTTAGCCAAGAGTCTCGTGGCGCGACGCTAG

Protein sequence: (SEQ ID NO: 224)

MERLLESIYINARPAMELRLSLTSSGRKRMVKIVDGEEVEVLPGEVQGILEAQKRDVGILADFLAKSLVARR

RS011

DNA sequence: (SEQ ID NO: 104)

ATGGAATGCCACGTTCGTCCCGCCACGAGCAGAGATGCAGCAGCGATAAGCTGCGTAGTTATAGCCGCCCTGCGTGAGTC
AAATTCACAGGACTATCCGCCTGATGTGATCGCTCAGGTTGAGCAGAGCTTTTCTCCTGAAGCCATCACCACACAGCTTA
CGAAGCGTAGGGTCTTCGTAGCCTTATTGGGCGAAAACATTATTGGCACTGCCGGTCTCGACGGTGACGTCGTCAGAAGT
GTTTTCGTTGACCCAGCTCACCAGAAAGGCGGTATCGGGCGGCATTTGATGGATGTCATTCATACAACTGCTGCCAGCGC
GGGAGTTGGAGCTGTACGTGTGCCATCGTCGATTACAGCTGAAAGGTTTTATACCGCATTGGGTTATCAGAAAATCCGCG
ACGAGTTTCATGGGGCGGAGCGCACCATCGTTATGGAGAAGCGGCTGTAG

Protein sequence: (SEQ ID NO: 225)

MECHVRPATSRDAAAISCVVIAALRESNSQDYPPDVIAQVEQSFSPEAITTQLTKRRVFVALLGENIIGTAGLDGDVVRS
VFVDPAHQKGGIGRHLMDVIHTTAASAGVGAVRVPSSITAERFYTALGYQKIRDEFHGAERTIVMEKRL

FIGURE 31D

RS12

DNA sequence: (SEQ ID NO: 105)

TTGTGGTTGACCTGCACGCCACAGCAGGATGTGCAGGCGGCGTTAGCTACAGCGTCGATACTCCTGGGCCAGTTCCACCA
GTTGGGCGTGCAGCTCGGTCGGTACACTAGCCTCGACCCGCTTGAGGAAGTCGAGAAGAACGCTTCTGCACTGCCGTCTC
CTGCTTGGAAAACGGATTCTACTAAGTTCAGCGTGGTACTGAAATCGGGGGGCAGGTCAATCGACAAAGGTATCCCGACC
GCAGGTTTGTTGGCCCACGTGATGGTGGCCAAGTTTGCCGATCACTTGCCGCTGTACCGGCAGGAGAAAATCTTTGGCCG
CGCCGGGCTGGCAATTGCTCGCTCGACCCTGGCGCAGTGGGTCGGACAAACCGGCGTGCGGCTTCAGCCACTGGTCGATG
CACTGCGTGAAGCCGTGCTGAACCAGGGCGTGATCCACGCTGATGAAACACCGGTGCAAATGCTTGCGCCAGGCGAGAAG
AAAACCCACCGGGCCTATGTCTGGGCGTACAGCACGACGCCGTTTTCAGGGCTCAAAGCGGTGGTTTACGACTTCAGCCC
AAGCCGTGCTGGCGAACATGCGCGCAACTTCCTGGGTGACTGGAACGGCAAGCTGGTCTGCGACGACTTCGCTGGCTACA
AAGCCGGTTTCGAACAAGGCATCACTGAAATCGGCTGCATGGCCCACGCCCGGCGCAAGTTCTTTGATTTGCACGTGGCG
AACAAAAGTCAGCTGGCTGAACAGGCCCTGCACTCGATCAGCGGCTTGTACGAGGTCGAACGTCAGGCGCGGGACATGAG
TGATGAAGAGCGCTGGCGAATACGACAAGAATTGGCGGTGCCGATCCTCAAAAAACTGCATGACTGGATGTTGGCTCAGC
GAGACCTGGTGCCCAATGGATCAGCCACGGCCAAAGCCCTCGATTACAGCCTGAAACGCTGGGTAGCGCTGACGCGCTAC
CTGGACGATGGGGCTGTGCCCATCGATAACAATCAGGTCGAGAACCAAATACGGCCATGGGCGCTCGGGCGTTCGAACTG
GCTGTTTGCCGGGTCGCTGCGCAGTGGTAAACGGGCGGCTGCAATCATGAGCCTGATCTAG

Protein sequence: (SEQ ID NO: 226)

LWLTCTPQQDVQAALATASILLGQFHQLGVQLGRYTSLDPLEEVEKNASALPSPAWKTDSTKFSVVLKSGGRSIDKGIPT
AGLLAHVMVAKFADHLPLYRQEKIFGRAGLAIARSTLAQWVGQTGVRLQPLVDALREAVLNQGVIHADETPVQMLAPGEK
KTHRAYVWAYSTTPFSGLKAVVYDFSPSRAGEHARNFLGDWNGKLVCDDFAGYKAGFEQGITEIGCMAHARRKFFDLHVA
NKSQLAEQALHSISGLYEVERQARDMSDEERWRIRQELAVPILKKLHDWMLAQRDLVPNGSATAKALDYSLKRWVALTRY
LDDGAVPIDNNQVENQIRPWALGRSNWLFAGSLRSGKRAAAIMSLI

RS13

DNA sequence: (SEQ ID NO: 106)

ATGGTGAGGCGGCGGAGGGTCGCGGTGGCGCGCGAATGCCTGAGCCTGTCGAGCGCACCGAACCAGGTCTTGTCGATGGA
TTTCGTCTTCGACGCGCTCAGCACTGGGCGACGGATCAAATGCCTGACGGTGGTCGATGACTTCACCAAGGTGTCGGTCG
ACATCTTGGTGGAGTACGGTATCAGCGGTTTTCGTGTCACGCGGGCGCTGGACGAGATGGCGCGGTTTCGTGGCTACCCG
CAGGCGATCCGCACCGACCAGGGCCCCGAGTTCACCGGCAAGGCGCTTGATCAGTGGGCCTGTCAGCGTGACATCAAGTT
GAAGCTGATTCAGCCTGGCCAGCCCACGCAGAGCGCCTTCATCGAGTCATTCAACGGCAAGTTCCGGGGCGAATGCCTCA
ATGAGCACTGCTCGCTGGTCGAAGCCAGAATCCGTATCGCGGCTTGGCGGGATTACAACGAGCACCGACCACACAGCGCC
ATTGGCAATCTCTCCCCGGCAGAGCTTGCTGCGAAGTGGCGAACCAACCAGCAGCAGCTGAAGCGGGAAAAGTTGATATC
AACCCCATAG

Protein sequence: (SEQ ID NO: 227)

MVRRRRVAVARECLSLSSAPNQVLSMDFVFDALSTGRRIKCLTVVDDFTKVSVDILVEYGISGFRVTRALDEMARFRGYP
QAIRTDQGPEFTGKALDQWACQRDIKLKLIQPGQPTQSAFIESFNGKFRGECLNEHCSLVEARIRIAAWRDYNEHRPHSA
IGNLSPAELAAKWRTNQQQLKREKLISTP

RS14

DNA sequence: (SEQ ID NO: 107)

ATGCATATCCAATCGTTGGGGGCTACTGCCTCCTCGCTGAATCAGGAGCCTGTCGAAACCCCGTCGCAGGCAGCGCATAA
GTCCGCCAGCTTGCGTCAGGAACCTTCAGGGCAAGGTCTCGGGGTTGCCCTAAAGAGCACGCCGGGAATACTTTCCGGGA
AGTTGCCGGAAAGCGTTAGCGACGTGCGTTTCAGCAGTCCCCAAGGGCAAGGGGAGTCCCGTACTCTGACTGACTCGGCA
GGGCCGCGGCAGATCACTCTGCGCCAGTTTGAGAACGGAGTCACCGAGCTACAGCTCAGTCGGCCACCATTGACCAGTCT
GGTCCTAAGCGGCGGTGGTGCCAAAGGTGCGGCATACCCGGGAGCAATGCTGGCGCTAGAAGAGAAAGGCATGCTCGATG
GCATCCGCAGCATGTCCGGTTCGTCCGCTGGCGGCATCACCGCCGCCCTTTTGGCCTCAGGTATGAGCCCGGCGGCGTTC
AAGACCCTTTCCGACAAGATGGATCTTATTTCGCTGCTCGACAGCTCGAACAAGAAGCTGAAGCTGTTCCAACACATTAG
CAGCGAGATCGGCGCATCGCTGAAAAAGGGCTTGGGCAACAAGATCGGCGGCTTCTCTGAGTTGCTGCTCAATGTACTCC
CACGCATAGATTCGCGGGCTGAGCCCCTAGAACGCCTATTGCGCGACGAGACACGCAAGGCCGTGCTCGGACAGATCGCT
ACGCATCCAGAGGTTGCACGCCAGCCGACCGTTGCCGCCATCGCCAGCAGATTGCAGTCCGGCTCCGGAGTCACCTTTGG
CGATCTAGATCGGTTGAGTGCTTACATTCCCCAGATTAAGACGCTGAACATCACAGGTACGGCCATGTTCGAGGGGCGTC
CGCAATTAGTGGTGTTCAATGCCAGCCACACACCGGATCTGGAGGTCGCCCAGGCGGCACATATCTCCGGTTCCTTCCCA
GGAGTGTTCCAGAAGGTCAGCTTGAGTGATCAGCCGTACCAGGCCGGCGTAGAGTGGACAGAATTCCAGGATGGCGGGGT
GATGATTAACGTGCCGGTCCCTGAGATGATCGACAAGAATTTTGACAGCGGGCCACTGCGGCGCAACGACAACCTGATCC
TTGAGTTCGAGGGCGAAGCTGGGGAGGTAGCGCCCGACCGAGGTACTAGGGGCGGCGCGCTCAAGGGCTGGGTCGTCGGG
GTGCCTGCCCTGCAGGCGCGCGAAATGCTGCAGCTCGAGGGCCTGGAGGAATTGCGCGAGCAAACCGTTGTGGTGCCGTT
GAAGAGCGAGCGCGGTGATTTCAGTGGCATGCTCGGTGGCACCTTGAACTTCACCATGCCGGACGAGATCAAGGCGCATC
TTCAGGAGCGCCTCCAGGAGCGAGTCGGTGAACATCTGGAGAAACGTCTTCAGGCTTCAGAGCGTCATACCTTCGCTTCT

FIGURE 31E

CTCGACGAGGCGCTGCTGGCACTTGATGACAGTATGCTCACCAGTGTTGCTCAACAGAACCCGGAGATCACAGACGGGGC
GGTGGCTTTTCGCCAGAAGGCGCGGGATGCGTTCACCGAGCTGACTGTCGCTATCGTTAGCGCCAATGGCTTGGCGGGTA
GGCTCAAGTTGGACGAGGCTATGCGCTCCGCTCTTCAGCGACTCGATGCGCTGGCAGATACTCCGGAACGCCTAGCATGG
TTGGCAGCTGAGTTGAACCATGCTGATAACGTTGATCATCAGCAGTTACTCGATGCCATGCGCGGGCAGACGGTGCAGTC
GCCGGTGCTCGCCGCTGCGTTAGCAGAGGCGCAGCGCCGCAAAGTGGCGGTTATTGCCGAGAACATTCGTAAGGAAGTTA
TCTTCCCCTCTCTGTATCGCCCTGGCCAGCCGGATTCCAACGTAGCTCTGTTACGTCGGGCGGAGGAGCAGCTACGGCAT
GCCACCAGTCCGGCGGAAATCAATCAAGCGCTGAACGATATCGTCGACAACTACTCGGCACGAGGCTTCCTGCGTTTCGG
CAAACCCTTGAGTTCGACTACCGTTGAGATGGCTAAGGCTTGGCGGAATAAGGAGTTCACATGATT

Protein sequence (SEQ ID NO:228)

MHIQSLGATASSLNQEPVETPSQAAHKSASLRQEPSGQGLGVALKSTPGILSGKLPESVSDVRFSSPQGQGESRTLTDSA
GPRQITLRQFENGVTELQLSRPPLTSLVLSGGGAKGAAYPGAMLALEEKGMLDGIRSMSGSSAGGITAALLASGMSPAAF
KTLSDKMDLISLLDSSNKKLKLFQHISSEIGASLKKGLGNKIGGFSELLLNVLPRIDSRAEPLERLLRDETRKAVLGQIA
THPEVARQPTVAAIASRLQSGSGVTFGDLDRLSAYIPQIKTLNITGTAMFEGRPQLVVFNASHTPDLEVAQAAHISGSFP
GVFQKVSLSDQPYQAGVEWTEFQDGGVMINVPVPEMIDKNFDSGPLRRNDNLILEFEGEAGEVAPDRGTRGGALKGWVVG
VPALQAREMLQLEGLEELREQTVVVPLKSERGDFSGMLGGTLNFTMPDEIKAHLQERLQERVGEHLEKRLQASERHTFAS
LDEALLALDDSMLTSVAQQNPEITDGAVAFRQKARDAFTELTVAIVSANGLAGRLKLDEAMRSALQRLDALADTPERLAW
LAAELNHADNVDHQQLLDAMRGQTVQSPVLAAALAEAQRRKVAVIAENIRKEVIFPSLYRPGQPDSNVALLRRAEEQLRH
ATSPAEINQALNDIVDNYSARGFLRFGKPLSSTTVEMAKAWRNKEFT

RS15

DNA sequence: (SEQ ID NO: 108)

ATGATTGATACATGGCTGGCACAGTGGGGCTTGAGACTTCCCTCGAGCAACGATGCCACGTTGCGGCTGCAACCGGCAGA
GGGACCGGAACTGGTTATGGAGCGCCTCGAGGGCGGTTGGCTTTTCGTCGTCGAGTTGGGACTTGTGCCTTCAGGGTTAC
CGCTGGGTGTGATCTTGCAATTGTTACAAGTGAACTCTCCATTCTCATCCTTGGCACCGGTGAAACTTGCGGCGGACGAT
GCCGGTAGACTTGTGCTCTGGGCTGAGGCACGTGATGGCGTTGACGATGTGGATGCACTGAACCGCTTGCACGATAGGCT
GCGGGAAGGACATTCACGATTAGTGCCATTGCTAGAGCCCACGGGTGAGTTGGTTCCAGCTCAGATACAAACCAGCGCGT
TAGTGTTCGTTTGA

Protein sequence: (SEQ ID NO: 229)

MIDTWLAQWGLRLPSSNDATLRLQPAEGPELVMERLEGGWLFVVELGLVPSGLPLGVILQLLQVNSPFSSLAPVKLAADD
AGRLVLWAEARDGVDDVDALNRLHDRLREGHSRLVPLLEPTGELVPAQIQTSALVFV

Figure 31

Big Island: Overall Nucleotide Homology

Total 84830bp

Bp #	Species, strain, gene name	Accession #	Evalue / %identity
23559-25465:	X. axonopodis pv. Citri strain 306	AE011864	83%

Only stretches within the 20437-25465bp are homologous to X. axonopodis pv. Citri strain 306. A total of 1060bp, not contiguous, from this region are homologous to X. axonopodis.

33872-38412:	P. aeruginosa, PA14, pvrR	AF482691	0.0; 99%
40989-46535:	P. aeruginosa, PA01, PA2128-2132	AE004640	80%

Only stretches within the 40989-46535bp region are homologous to PA01. A total of 2406bp, not contiguous, from this region are homologous to PA01.

48266-49533:	P. putida, plasmid pWWO	AJ344068	96%

Only stretches within the 48266-49533bp are homologous to P. putida, plasmid pWWO. A total of 780bp, not contiguous, from this region are homologous to P. putida.

56824-58706:	P. syringae pv. maculicola, plasmid pFKN	AF359557	83%

Only stretches within the 56824-58706bp are homologous to P. syringae pv. maculicola, plasmid pFKN. A total of 1882bp, not contiguous, from this region are homologous to P. syringae.

64748-64942:	P. aeruginosa, PA103, exoU,	U97065	1E-85/96%
82447-85179:	P. aeruginosa, PA01, PA0984-0985	AE004531	0.0/97%
85334-855542:			3E-80/94%
93200-93317:	P. aeruginosa, PA158	X73064	7E-50/98%
108075-108610:	P. aeruginosa, SG17M, plasmid pKLC102	AF285416	0.0/91%
100119-101054:	P. aeruginosa, PA01, PA3849,	AE4802	0.0/98%

Figure 32

Small Island: Overall Nucleotide Homology

Total 10848bp

Bp #	(Species, strain, gene name	Accession #)	E value / %identity
877-1632:	*P. aeruginosa*, PA01, PA0977	AE004531	0.0/94%
877-2603:	*P. aeruginosa*, SG17M plasmid pKLC102	AF285416	0.0/92%
2978-6471:	*P. aeruginosa*, PA01, PA0978-81	AE004531	0.0/99%
7035-7999:	P. syringae pv. maculicola, plasmid pFKN	AF359557	83%

Only stretches within the 7035-7999 bp are homologous to P. syringae pv. maculicola, plasmid pFKN. A total of 534bp, not contiguous, from this region are homologous to P. syringae.

7999-8284	*P. aeruginosa*, PA103, exoU,	AF27291	E-136/96
8000-8080:	*P. aeruginosa*, PA01 intragenic region	AE004531	E-18/91%
8120-8259:	*P. aeruginosa*, PA01 intragenic region	AE004531	E-24/85%
8272-8860:	*P. aeruginosa*, PA01 intragenic region	AE004531	E-176/88%
8470-11724	*P. aeruginosa*, PA103, exoU,	AF27291	0.0/99%

Figure 33

FIGURE 34A

ORF ID	Strand	Left end	Right end	ORF length (aa)	G+C content (%)	Location prediction	Gene name	Gene function	Protein with the highest identity (gene name / species strain)	E-value (% identity)	Proteins with lesser identity Cut off 30%	GenBank accession no.
		725	800		51				tRNA-Lys / *P. aeruginosa* PAO1			AE004868
RL001	-	878	1,202	pseudogene	56			Hypothetical protein	PA0977 / *P. aeruginosa* PAO1	3E-45 (91)		AAG04366
RL002		1,169	2,452	427	60	cytoplasm	*xerC*	Integrase	Integrase / *P. aeruginosa* SG17M	0.0 (91)	STY4666	AAG02084
RL003	-	2,449	4,368	639	58	cytoplasm		Hypothetical protein	PSPTO0831 / *P. syringae* pv. *tomato* DC3000	E-180 (52)	XAC2196,XCC3121, STY4665	AAO54371
RL004		4,903	5,436	177	56	cytoplasm	*dcd-2*	Deoxycytidine deaminase	MK1566 / *Methanopyrus kandleri* AV19	1E-13 (33)		AAM02709
RL005	+	5,436	6,146	236	51	inner membrane		Hypothetical protein	No significant similarity			
RL006		6,143	6,682	179	45	cytoplasm	*dut*	Deoxycytidine triphosphate deaminase	VNG0245G / *Halobacterium* sp. NRC-1	1E-07 (32)		AAG19843
RL007	-	6,887	7,402	171	48	outer membrane and periplasm		Hypothetical protein	No significant similarity			
RL008		7,829	9,760	643	60	cytoplasm		Hypothetical protein	Protein fused from putative helicase (*Methanosarcina acetivorans* C2A) and hypothetical protein PA1935 (*P. aeruginosa* PAO1)	7E-09 (25) / 8E-16 (32)		AAM05583 / AAG05328
RL009	-	9,757	12,180	807	60	inner membrane		Hypothetical protein	PA1939 / *P. aeruginosa* PAO1	2E-69 (30)		AAG05327
RL010		12,358	12,660	100	54	inner membrane		Hypothetical protein	No significant similarity			
RL011	-	13,102	13,452	116	60	cytoplasm	*parE*	Plasmid stabilization protein parE	AGR_C_2415p / *Agrobacterium tumefaciens* C58	8E-22 (45)		AAK87104
RL012		13,456	13,728	90	62	cytoplasm		Putative transcription regulator	AGR_C_2413p / *Agrobacterium tumefaciens* C58	2E-13 (41)	STY3093	AAK87103
RL013	+	13,847	14,191	114	48	inner membrane		Hypothetical protein	No significant similarity			
RL014		14,217	15,728	503	58	inner membrane		Hypothetical protein	XAC2186 / *Xanthomonas axonopodis* pv. *citri* 306	1E-175 (65)	XCC3117, STY4579	AAM37039
RL015	-	15,725	16,066	113	62	inner membrane		Hypothetical protein	No significant similarity			
RL016		16,066	17,448	460	64	outer membrane and periplasm		Hypothetical protein	PSPTO0848 / *P. syringae* pv. *tomato* DC3000	1E-138 (57)	XCC3116, STY4577	AAO54383

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF
Applicants: Laurence Rahme et al.
Filing Date: September 12, 2003 Serial No.: Not Yet Assigned
Page 75 of 118 Customer No.: 21559

FIGURE 34B

RL017	-	17,466	18,404	312	65	outer membrane and periplasm		Hypothetical protein	PSPO0849 / *P. syringae* pv. *tomato* DC3000	E-125 (72)	XAC2283, STY4576	AAO54384
RL018		18,404	18,885	141	62	outer membrane and periplasm		Hypothetical protein	SG52 // *P.aeruginosa* SG17M	3E-22 (49)	STY4575	AAM87185
RL019	+	19,044	19,262	72	55	cytoplasm		Hypothetical protein	No significant similarity		XAC2275	
RL020		19,259	19,918	219	62	outer membrane and periplasm	*dsbG*	Putative protein disulfide isomerase	PA0982 // *P.aeruginosa* PAO1	1E-92 (93)		AAG04971
RL021	-	19,915	20,199	94	57	cytoplasm		Hypothetical protein	PSPTO0858 / *P. syringae* pv. *tomato* DC3000	2E-23 (55)	STY4573, STY4572	AAO54393
RL022		20,196	23,138	980	64	cytoplasm		Hypothetical protein	PSPTO0859 // *P. syringae* pv. *tomato* DC3000	0.0 (64)		AAO54394
RL023	-	23,138	23,581	147	64	inner membrane		Hypothetical protein	PSPTO0860 / *P. syringae* pv. *tomato* DC3000	2E-48 (74)		AAO54395
RL024		23,559	25,064	501	63	outer membrane and periplasm		Hypothetical protein	PSPTO0861 // *P. syringae* pv. *tomato* DC3000	E-148 (54)	STY4570, ORF1094	AAO54396
RL025	-	25,048	25,932	294	66	outer membrane and periplasm		Hypothetical protein	PSPTO0862 / P. *syringae* pv. *tomato* DC3000	1E-90 (62)	XAC2272, STY4569	AAO54397
RL026		25,929	26,588	219	60	inner membrane		Hypothetical protein	PSPTO0863 // *P.syringae* pv. *tomato* DC3000	1E-71 (59)	STY4568	AAO54398
RL027	-	26,585	26,971	128	65	inner membrane		Hypothetical protein	XAC2271 / *X. axonopodis* pv. *citri* 306	4E-29 (50)		AAM37124
RL028		26,982	27,338	118	60	inner membrane		Hypothetical protein	ORF1116 // *P. putida* (plasmid pWW0)	6E-23 (50)	XAC2270, STY4566	CAC86817
RL029	-	27,356	27,595	79	63	inner membrane		Hypothetical protein	C54 / *P. aeruginosa* C	1E-09 (38)	STY4565	AAN62148
RL030		27,592	27,951	119	66	inner membrane		Putative type III effector Hop protein	PSPTO0869 // *P.syringae* pv. *tomato* DC3000	4E-25 (53)	XAC2270, STY4566	AAO54404
RL031	-	28,024	28,329	101	57	cytoplasm		Hypothetical protein	No significant similarity		PA4736, PA4737	
RL032		28,502	28,813	103	49	outer membrane and periplasm		Hypothetical protein	PA0713 // *P. aeruginosa* PAO1	9E-16 (43)		AAG04102
RL033	-	28,810	29,967	385	47	cytoplasm		Hypothetical protein	No significant similarity			
RL034		30,098	31,579	493	59	inner membrane		Putative DNA helicase	PSPTO0879 // *P.syringae* pv. *tomato* DC3000	0.0 (65)		AAO54413

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF
Applicants: Laurence Rahme et al.
Filing Date: September 12, 2003 Serial No.: Not Yet Assigned
Page 76 of 118 Customer No.: 21559

FIGURE 34C

RL035	-	31,590	32,240	216	62	inner membrane		Hypothetical protein	PSPTO0880 / *P. syringae* pv. *tomato* DC3000	2E-74 (63)	XAC2260, STY4563	AAO54414
RL036	+	32,572	35,367	931	60	inner membrane		sensor of two-component regulatory system	ORF4 / *P. aeruginosa* PA14 / / RcsC / *S. enterica* subsp. *enterica* ser. Typhi CT18	0.0 (910) / 8E-98 (32)	YfES?	AAM15532 / GAD07502
RL037	+	35,364	36,563	399	59	cytoplasm	*pvrR*	regulator of two-component regulatory system; adhesion and antibiotic resistance	PvrR / *P. aeruginosa* PA14 / / VieA / *X. campestris* pv. *campestris* ATCC 33913	0.0 (100) / 2E-55 (34)	pvrR	AAM15533 / AAM41975
RL038	+	36,644	39,898	1,084	63	inner membrane	*rcsC*	sensor of two-component regulatory system	STM2271 / *S. Typhimurium* LT2	4E-85 (32)		AAL21172
RL039	+	39,898	40,593	231	59	cytoplasm	*rcsB*	regulator of two-component regulatory system	Z3476 / *E. coli* O157:H7 EDL933	8E-33 (39)		AAG57352
RL040	-	40,687	41,353	298	67	outer membrane and periplasm	*cupD5*	Probable pili assembly chaperone / adhesion and protein secretion	PA2132 / *P. aeruginosa* PAO1	4E-65 (62)		AAG05520
RL041	-	41,343	42,689	448	63	outer membrane and periplasm	*cupD4*	adhesion and protein secretion	PA2131 / *P. aeruginosa* PAO1	0.0 (70)	STY0370	AAG05519
RL042	-	42,686	45,298	1870	66	outer membrane and periplasm	*cupD3*	Probable fimbrial biogenesis usher / adhesion and protein secretion	PA2130 / *P. aeruginosa* PAO1	0.0 (67)	STY0371	AAG05518
RL043	-	45,282	46,028	248	65	outer membrane and periplasm	*cupD2*	Probable pili assembly chaperone / adhesion and protein secretion	PA2129 / *P. aeruginosa* PAO1	1E-83 (64)	STY0372	AAG05517
RL044	-	46,107	46,665	182	65	outer membrane and periplasm	*cupD1*	Probable fimbrial precursor / adhesion and protein secretion	PA2128 / *P. aeruginosa* PAO1	6E-66 (72)	STY0373	AAG05516
RL045	-	46,836	47,101	pseudogene	55			Recombination	Transposase / *E. coli* (plasmid p1658/97)	2E-36 (82)		AAO49572
RL046	-	47,103	47,849	248	61	inner membrane		Hypothetical protein	PSPTO0880 / *P. syringae* pv. *tomato* DC3000	E-83 (59)	XAC2260, STY4563	AAO54414
RL047	-	47,849	50,083	744	65	inner membrane		Hypothetical protein	SG68 / *P. aeruginosa* SG17M	0.0 (66)	XAC2259, STY4562	AAN62290
RL048	-	50,087	50,844	85	59	cytoplasm		Hypothetical protein	No significant similarity			

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF
Applicants: Laurence Rahme et al.
Filing Date: September 12, 2003 Serial No.: Not Yet Assigned
Page 77 of 118 Customer No.: 21559

FIGURE 34D

RL049	-	50,353	50,853	166	66	outer membrane and periplasm		Hypothetical protein	SG69 / *P. aeruginosa* SG17M	7E-34 (51)	STY4560	AAN62291
RL050	-	50,860	51,491	193	64	outer membrane and periplasm		Hypothetical protein	SG70 / *P. aeruginosa* SG17M	5E-38 (46)	STY4559	AAN62292
RL051	-	51,416	52,171	251	65	outer membrane and periplasm		Hypothetical protein	C68 / *P. aeruginosa* C	5E-50 (52)	STY4558	AAN62162
RL052	-	52,182	52,871	229	64	inner membrane		Hypothetical protein	C69 / *P. aeruginosa* C	5E-05 (25)		AAN62163
RL053	-	53,019	53,235	pseudogene	61			Recombination	Transposase / *P. putida* (plasmid pWW53)	7E-09 (54)		BAB59051
RL054	-	53,482	54,513	334	47	cytoplasm		Hypothetical protein	PA2223 / *P. aeruginosa* PAO1	2E-56 (45)		AAG05611
RL055	+	54,524	55,189	221	50	cytoplasm		Hypothetical protein	PA2222 / *P. aeruginosa* PAO1	2E-56 (48)		AAG05610
RL056	+	55,272	55,946	224	49	cytoplasm		Hypothetical protein	PA2224 / *P. aeruginosa* PAO1	2E-06 (23)	XAC4135	AAG05612
RL057	+	56,030	56,275	81	59	cytoplasm		Hypothetical protein	No significant similarity			
RL058	-	56,348	56,707	119	48	cytoplasm		Hypothetical protein	CAC2557 / *Clostridium acetobutylicum*	2E-13 (34)	XCC0996	AAK80508
RL059	-	56,768	58,303	511	59	cytoplasm		Transposase	PP4439 / ISPpu14 ORF3 / *P. putida* KT2440	0.0 (90)		AAN70015
RL060	-	58,367	58,702	111	63	cytoplasm		Transposase	PP4438 / ISPpu14 ORF2 / *P. putida* KT2440	2E-59 (98)		AAN70014
RL061	-	58,753	59,019	88	57	cytoplasm		Transposase	PP4437 / ISPpu14 ORF1 / *P. putida* KT2440	7E-39 (88)		AAN70013
RL062	-	59,276	60,502	408	46	cytoplasm		Hypothetical protein	No significant similarity			
RL063	-	60,826	63,075	749	63	inner membrane		Plasmid-related protein	XAC2243 / *X. axonopodis* pv. *citri* 306	0.0 (69)		AAM37096
RL064	-	63,180	64,631	483	64	cytoplasm		Plasmid-related protein	XAC2242 / *X. axonopodis* pv. *citri* 306	1E-133 (65)		AAM37095
RL065	-	64,661	65,266	201	63	cytoplasm		Hypothetical protein	XAC2241 / *X. axonopodis* pv. *citri* 306	1E-54 (56)		AAM37094
RL066	-	65,358	65,612	84	60	cytoplasm		Hypothetical protein	XF1757 / *X. fastidiosa* 9a5c	9E-09 (40)	XAC2240	AAF84566
RL067	-	65,680	66,042	120	58	cytoplasm		Hypothetical protein	C77 / *P. aeruginosa* C	8E-23 (42)	XAC2239	AAN62171
RL068	-	66,112	66,387	91	63	cytoplasm		Hypothetical protein	No significant similarity			
RL069	-	66,384	67,073	229	60	inner membrane		Hypothetical protein	XF1760 / *X. fastidiosa* 9a5c	3E-50 (53)	XAC2237	AAF84569

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF
Applicants: Laurence Rahme et al.
Filing Date: September 12, 2003 Serial No.: Not Yet Assigned
Page 78 of 118 Customer No.: 21559

FIGURE 34E

RL070	-	67,077	67,427	116	62	cytoplasm		Hypothetical protein	No significant similarity			
RL071	-	67,663	68,370	235	53	cytoplasm		Hypothetical protein	SG91 / *P. aeruginosa* SG17M	6E-35 (40)		AAN62312
RL072	-	68,855	69,091	78	52	cytoplasm		Hypothetical protein	STY4535 / *S. enterica* subsp. *enterica* ser. Typhi CT18	7E-11 (64)	XCC1772, XAC2217	CAD09214
RL073	+	69,111	69,377	88	57	cytoplasm		Hypothetical protein	No significant similarity			
RL074	+	69,426	69,965	179	57	cytoplasm		Hypothetical protein	XF1761 / *X. fastidiosa* 9a5c	1E-28 (62)	XAC2236, STY4584, XCC2093	AAM84570
RL075	-	70,626	71,192	188	58	cytoplasm		Hypothetical protein	No significant similarity			
RL076	+	71,191	71,835	214	60	cytoplasm		Hypothetical protein	No significant similarity			
RL077	-	72,107	72,544	145	68	outer membrane and periplasm	*pilM2*	Type IV B pilus / adhesion and and protein secretion	PilM / *S. typhimurium* (plasmid R64)	5E-04 (21)	PA4199, STY4540	BAB91693
RL078	-	72,573	73,901	442	63	outer membrane and periplasm	*pilV2*	Type IV B pilus / adhesion and and protein secretion	shufflon A / *E. coli* (plasmid R64)	3E-65 (38)	STY4550, XAC2151	C26421
RL079	-	73,906	74,847	313	65	inner membrane	*pilT2*	Type IV B pilus / putative peptidase / adhesion and and protein secretion	BfpF / *E. coli* (plasmid pB171)	8E-25 (27)	XAC2923, XCC2754, PA0396, PA0395	BAA84845
RL080	-	74,844	75,374	176	60	outer membrane and periplasm	*pilS2*	Type IV B pilus / adhesion and and protein secretion	PilS / *Shigella sonnei* P9 (plasmid ColIb-P9)	5E-15 (29)	STY4547	BAA75180
RL081	-	75,396	76,475	359	55	inner membrane	*pilR2*	Type IV B pilus / adhesion and and protein secretion	PilR / *S. typhimurium* (plasmid R64)	4E-51 (33)	BfpE, STY4546, XAC0697, XCC3423 (Type II), PA2676, PA3102, STY0164	BAB91688
RL082	-	76,475	78,055	526	63	inner membrane	*pilQ2*	ATPase / Type IV B pilus / adhesion and and protein secretion	PilQ / *S. typhimurium* (plasmid R64)		STY4545, bfpD, gspE (Type II)	BAB91689
RL083	-	78,064	78,597	177	69	outer membrane and periplasm	*pilP2*	Type IV B pilus / adhesion and and protein secretion	PilP / *S. typhimurium* (plasmid R64)	1E-06 (34)	STY4544	BAB91690
RL084	-	78,587	79,912	441	64	inner membrane	*pilO2*	Type IV B pilus / adhesion and and protein secretion	PilO / *S. enterica* subsp. *enterica* ser. Dublin	3E-15 (21)	STY4543	AAR81211

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF
Applicants: Laurence Rahme et al.
Filing Date: September 12, 2003 Serial No.: Not Yet Assigned
Page 79 of 118 Customer No.: 21559

FIGURE 34F

RL085	-	79,916	81,538	540	63	outmembrance	*pilN2*	Secretin / Type IV B pilus / adhesion and and protein secretion	PilN / *E. coli* K-12 (plasmid R721)	6E-84 (35)	bfpB, PA1382 (Type II), XF1527, pefD (general secretion pathway protein)	BAB12647
RL086	-	81,627	82,751	374	66	outer membrane and periplasm	*pilN2*	Type IV B pilus / adhesion and and protein secretion	STY4539 / *S. enterica* subsp. *enterica* ser. Typhi CT18	1E-48 (34)		CAD09317
RL087	+	83,023	83,349	108	40	inner membrane		Colicin immunity protein	PA0984 / *P. aeruginosa* PAO1	4E-55 (94)		AAG04373
RL088	+	83,374	84,870	499	47	inner membrane		Colicin-like toxin (pyocin S5)	PA0985 / *P. aeruginosa* PAO1	0.0 (97)	PA1150	AAG04374
RL089	-	85,558	87,531	657	64	inner membrane		DNA Helicase	DNA helicase / *Dictyostelium discoideum* AX14	4E-28 (24)	PA0799	AAO51158
RL090	-	87,528	89,417	629	60	cytoplasm		Hypothetical protein	No significant similarity			
RL091	-	89,551	90,021	156	44	cytoplasm		Similar to luminal binding protein	Riorf142 / *Rhizobium rhizogenes* MAFF03-01724	5E-29 (41)		BAB16261
RL092	-	90,095	92,020	641	62	cytoplasm	*topA*	Topoisomerase I	XFa0003 / *X. fastidiosa* 9a5c (plasmid pXF51)	1E-151 (43)	PA5011, XF1776, XF0920, STM1298, XCC3755, STY1336	AAF85572
RL093	+	92,340	92,573	77	58	cytoplasm		Hypothetical protein	No significant similarity			
RL094	-	92,712	92,957	81	59	cytoplasm		Hypothetical protein	No significant similarity			
RL095	-	93,421	93,891	156	61	cytoplasm	*ssb*	Single-stranded DNA binding protein	C102 / *P. aeruginosa* C	6E-32 (53)	XAC2211, PA4232	AAN62318
RL096	-	93,905	94,498	177	60	cytoplasm		Hypothetical protein	XF1779 / *X. fastidiosa* 9a5c	8E-39 (53)	XAC2210	AAF84387
RL097	-	94,444	95,172	242	63	cytoplasm		Hypothetical protein	ORF6 / *Pseudomonas* sp. B13	7E-66 (52)	XAC2209, STY4529	CAD60670
RL098	-	95,442	95,681	79	46	inner membrane		Hypothetical protein	No significant similarity			
RL099	-	95,672	95,947	91	60	cytoplasm		Hypothetical protein	No significant similarity			
RL100	-	95,944	97,269	441	60	cytoplasm		Hypothetical protein	ORF5 / *Pseudomonas* sp. B13	1E-68 (39)	XAC2208, STY4528	CAD60669
RL101	-	97,266	98,033	255	58	cytoplasm		Hypothetical protein	ORF4 / *Pseudomonas* sp. B13	2E-33 (40)	XF1782, STY4526	CAD60668
RL102	-	98,061	99,800	579	58	cytoplasm		Hypothetical protein	SG102 / *P. aeruginosa* SG17M	1E-116 (49)	XF1783, XF1784, STY4523	AAN62323
RL103	-	99,797	100,051	84	60	cytoplasm		Hypothetical protein	No significant similarity			
RL104	-	100,048	101,064	1338	64	cytoplasm		Hypothetical protein	PA3849 / *P. aeruginosa* PAO1	0.0 (96)	STY2463, STY1607, STM2226	AAG07286
RL105	-	101,064	101,297	77	65	cytoplasm		Hypothetical protein	No significant similarity			

FIGURE 34G

RL106		101,290	101,784	164	62	cytoplasm		Hypothetical protein	No significant similarity			
RL107	-	101,777	102,034	85	60	cytoplasm		Hypothetical protein	No significant similarity			
RL108		102,031	102,558	175	62	cytoplasm		Hypothetical protein	No significant similarity			
RL109	-	102,745	104,107	pseudogene	61		*dnaB*	DNA replication and recombination	Replicative DNA helicase DnaB / *Pseudomonas* sp. SLT2001 (plasmid pQBR55)	E-130 (60)	PA4931, STY4442, XAC1477, XF0361, XCC1434	CAD13464
RL110	-	104,282	104,989	235	60	cytoplasm		Putative phage protein	FNY0875/ *Fusobacterium nucleatum* subsp. *vincentii* 49256	2E-21 (34)		BAA24090
RL111	-	104,986	105,687	233	64	cytoplasm		Hypothetical protein	STY1595 / *S .enterica* subsp. *enterica* ser. Typhi CT18	1E-16 (32)		CAD01841
RL112	-	105,687	106,373	228	64	inner membrane		Hypothetical protein	No significant similarity			
RL113	-	106,516	107,013	165	60	outer membrane and periplasm		Hypothetical protein	PA2226 / *P. aeruginosa* PAO1	2E-32 (47)		AAG05614)
RL114	-	107,010	107,690	226	58	cytoplasm		Hypothetical protein	ORF50 / *Pseudida* plasmid pDTG1	2E-25 (35)		AF491307
RL115	-	107,687	108,616	309	60	inner membrane	*soj*	Chromosome partitioning	Soj / *P. aeruginosa* SG17M	1E-84 (90)	XF1785, XAC2205, STY4521	AAG02083
		108,700	108,757		47				*attR* / *P. aeruginosa* PAO1			

Alignment: No_2 - embl|AL039136|HSM003612

Homo sapiens mRNA; EST DKFZp566K094_r1 (from clone DKFZp566)

```
Q:      18 DQTCDNLSQNPPHHLLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH SEQ ID NO: 230
           DQTCDNLSQNPPHHLLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH SEQ ID NO: 231
H:      29 DQTCDNLSQNPPHHLLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH SEQ ID NO: 232

           HYCCLWPVWMLCYMSW      93
           HYCCLWPVWMLCYMSW
           HYCCLWPVWMLCYMSW     256
```

Figure 35

Alignment: No_8 - embl|M79137|HSXT01285

EST01285 Subtracted Hippocampus, Stratagene (cat. #936205) H

```
Q:      18 QVQHPPLCLLDQHQQECIPPCLPPDHLQDPQHPFLLPDHHVPHLVVLIQPQLCRALAP SEQ ID NO: 233
75
           QVQHP .CLLDQHQQECIPPCLPPDHLQDPQHPFLLPDHHVP.LVVLIQPQLCRALAP SEQ ID NO: 234
H:      43 QVQHPXXCLLDQHQQECIPPCLPPDHLQDPQHPFLLPDHHVPXLVVLIQPQLCRALAP SEQ ID NO: 235
216
```

Figure 36

Alignment: No47 - swissnew|P35555|FBN1_HUMAN

FIBRILLIN 1 PRECURSOR.//:swiss|P35555|FBN1_HUMAN FIBRILLIN 1 PRECURSOR.//:trembl|L13923|HSFIBRLLN_1 product: "fibrillin"; Homo sapiens fibrillin mRNA, complete cds. //:gp|L13923|306746 product: "fibrillin"; Homo sapiens fibrillin mRNA, complete cds.

```
Q:      18 CGGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:236
           CGGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:237
H:    2617 CGGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:238

           GYFRIGQGHCVSGMGMGRGNPEPPVSGEMDDNSLSPEACYECKINGYPKRGRKRRSTNET SEQ ID NO:236
           GYFRIGQGHCVSGMGMGRGNPEPPVSGEMDDNSLSPEACYECKINGYPKRGRKRRSTNET SEQ ID NO:237
           GYFRIGQGHCVSGMGMGRGNPEPPVSGEMDDNSLSPEACYECKINGYPKRGRKRRSTNET SEQ ID NO:238

           DASNIEDQSETEANVSLASWDVEKTAIFAFNISHV-NKVRIL     178   SEQ ID NO:236
           DASNIEDQSETEANVSLASWDVEKTAIFAFNISHV NKVRIL           SEQ ID NO:237
           DASNIEDQSETEANVSLASWDVEKTAIFAFNISHVSNKVRIL    2778   SEQ ID NO:238
```

Figure 37

Alignment: No56 - trembl|AF088916|AF088916_1

gene: "EMI"; product: "elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.
//:trembl|AF088916|AF088916_1 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR. //:gp|AF088916|5353510 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR.
//:gpnew|AF162780|6693840 gene: "EMI"; product: "elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

```
Q:      7 DGDVYNPSTGVFTAPYDGRYLITATLTPERDAYVEAVLSVSNASVAQLHTAGYRREFLEY SEQ ID NO: 239
          DG..Y:P.TGVFTAP. GRYL::A.LT .R.. VEAVLS SN..VA::.:.GY. E LE  SEQ ID NO: 240
H:    896 DGGYYDPETGVFTAPLAGRYLLSAVLTGHRHEKVEAVLSRSNQGVARVDSGGYEPEGLE- SEQ ID NO: 241

          HRPPGALHTCGGP-GAFHLIVHLKAGDAV      94 SEQ ID NO: 239
          ::P .. :.. G. G.F.LI: L:AGD.V        SEQ ID NO: 240
          NKPVAESQPSPGTLGVFSLILPLQAGDTV     983 SEQ ID NO: 241
```

gene: "EMI"; product: "elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.
//:trembl|AF088916|AF088916_1 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR. //:gp|AF088916|5353510 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR.
//:gpnew|AF162780|6693840 gene: "EMI"; product: "elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

```
Q:      7 DGDVYNPSTGVFTAPYDGRYLITATLTPERDAYVEAVLSVSNASVAQLHTAGYRREFLEY SEQ ID NO: 242
          DG..Y:P.TGVFTAP. GRYL::A.LT .R.. VEAVLS SN..VA::.:.GY. E LE  SEQ ID NO: 243
H:    896 DGGYYDPETGVFTAPLAGRYLLSAVLTGHRHEKVEAVLSRSNQGVARVDSGGYEPEGLEN SEQ ID NO: 244

          HRPPGALHTCGGPGAFHLIVHLKAGDAV       94   SEQ ID NO: 242
          .  . :  : G  G.F.LI: L:AGD.V           SEQ ID NO: 243
          KPVAESQPSPGTLGVFSLILPLQAGDTV      983   SEQ ID NO: 244
```

Figure 38

Alignment: No59 - pironly|A35763|A35763

unnamed ORF; P.lividus 2-alpha collagen (COLL2-alpha) mRNA, complete cds.
//:pironly|A35763|A35763 collagen alpha 2 chain - sea urchin (Paracentrotus lividus)
(fragment)//:gp|J05422|159962 unnamed ORF; P.lividus 2-alpha collagen (COLL2-alpha)
mRNA, complete cds.

```
Q:     92 GENGSSGSQAPLQGLRGIFGLWGRRSRARFCGPR-PVARLGGGTSAGRELGL    142 SEQ ID NO: 245
          GE G.SG...P QG:RGI G: G...:.   GPR P .  GGG S G.. GL        SEQ ID NO: 246
H:    718 GEPGPSGENGP-QGVRGIPGVVGENGKTGRGGPRGPPGLRGGGGSRGERGGL    768 SEQ ID NO: 247
```

unnamed ORF; P.lividus 2-alpha collagen (COLL2-alpha) mRNA, complete cds.
//:pironly|A35763|A35763 collagen alpha 2 chain - sea urchin (Paracentrotus lividus)
(fragment)//:gp|J05422|159962 unnamed ORF; P.lividus 2-alpha collagen (COLL2-alpha)
mRNA, complete cds.

```
Q:     92 GENGSSGSQAPLQGLRGIFGLWGRRSRARFCGPR-PVARLGGGTSAGRELGL    142 SEQ ID NO: 248
          GE G.SG...P QG:RGI G: G...:.   GPR P .  GGG S.G.. GL        SEQ ID NO: 249
H:    718 GEPGPSGENGP-QGVRGIPGVVGENGKTGRGGPRGPPGLRGGGGSRGERGGL    768 SEQ ID NO: 250
```

Figure 39

Alignment: No60/63 - swiss|P20062|TCO2_HUMAN

TRANSCOBALAMIN II PRECURSOR.//:trembl|M60396|HSTCII_1 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds. //:gp|M60396|339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.

```
Q:       8 VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHFGN SEQ ID NO: 251
           VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHFGN SEQ ID NO: 252
H:     183 VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHFGN SEQ ID NO: 253

            VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKTYI SEQ ID NO: 251
            VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKTYI SEQ ID NO: 252
            VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKTYI SEQ ID NO: 253

            DLIFPDCLAPRVMLEPAA      145  SEQ ID NO: 251
            DLIFPDCLAPRVMLEPAA           SEQ ID NO: 252
            DLIFPDCLAPRVMLEPAA      320  SEQ ID NO: 253
```

TRANSCOBALAMIN II PRECURSOR.//:trembl|M60396|HSTCII_1 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds. //:gp|M60396|339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.

```
Q:       8 VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMA      47 SEQ ID NO: 254
           VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMA         SEQ ID NO: 255
H:     183 VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMA     222 SEQ ID NO: 256
```

TRANSCOBALAMIN II PRECURSOR.//:trembl|M60396|HSTCII_1 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds. //:gp|M60396|339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.

```
Q:       8 VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMA      47  SEQ ID NO: 257
           VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMA          SEQ ID NO: 258
H:     183 VEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMA     222  SEQ ID NO: 259
```

Figure 40

Alignment: No65 - swissnew|P23142|FBL1_HUMAN

FIBULIN-1 PRECURSOR.//:swiss|P37888|FBLD_HUMAN FIBULIN-1, ISOFORM D PRECURSOR.//:trembl|U01244|HS2444_1 product: "fibulin-1D"; Homo sapiens fibulin-1D mRNA, complete cds. //:gp|U01244|1621019 product: "fibulin-1D"; Homo sapiens fibulin-1D mRNA, complete cds.

```
Q:      18 RNCQDIDECVTGIHNCSINETCFNIQGGFRCLAFECPENYRRSAATLQQEKTDTVRCIKS SEQ ID NO: 260
           RNCQDIDECVTGIHNCSINETCFNIQG.FRCLAFECPENYRRSAATLQQEKTDTVRCIKS SEQ ID NO: 261
H:     521 RNCQDIDECVTGIHNCSINETCFNIQGAFRCLAFECPENYRRSAATLQQEKTDTVRCIKS SEQ ID NO: 262

           CRPNDVTCVFDPVHTISHTVISLPTFREFTRPEEIIFLRAITPPHPASQANIIFDITEGN SEQ ID NO: 260
           CRPNDVTCVFDPVHTISHTVISLPTFREFTRPEEIIFLRAITPPHPASQANIIFDITEGN SEQ ID NO: 261
           CRPNDVTCVFDPVHTISHTVISLPTFREFTRPEEIIFLRAITPPHPASQANIIFDITEGN SEQ ID NO: 262

           LRDSFDIIKRYMDGMTVGIRR     158  SEQ ID NO: 260
           LRDSFDIIKRYMDGMTVG: R          SEQ ID NO: 261
           LRDSFDIIKRYMDGMTVGVVR     661  SEQ ID NO: 262
```

Figure 41

Alignment: 80 - trembl|AF045447|AF045447_1

gene: "DPC4"; product: "deleted in pancreatic carcinoma"; Homo sapiens deleted in pancreatic carcinoma (DPC4) gene, exon 11 partial sequence and complete cds. //:trembl|U44378|HS443781_1 gene: "DPC4"; product: "Dpc4"; Human homozygous deletion target in pancreatic carcinoma (DPC4) mRNA, complete cds. //:pironly|S71811|S71811 probable transcription regulator MAD-4 - human//:gp|AF045447|2865657 gene: "DPC4"; product: "deleted in pancreatic carcinoma"; Homo sapiens deleted in pancreatic carcinoma (DPC4) gene, exon 11 partial sequence and complete cds. //:gp|U44378|1163234 gene: "DPC4"; product: "Dpc4"; Human homozygous deletion target in pancreatic carcinoma (DPC4) mRNA, complete cds.

```
Q:       6 PGSRIRGRVDTLQXNAPXXMMVKDEYVHDFEGQPXLXTEGHXIQTIQHPPXNRAXTETYX SEQ ID NO: 263
            PG  :.G   TLQ.NAP..MMVKDEYVHDFEGQP.L.TEGH.IQTIQHPP.NRA.TETY. SEQ ID NO: 264
H:     139 PGIDLSGL--TLQSNAPSSMMVKDEYVHDFEGQPSLSTEGHSIQTIQHPPSNRASTETYS SEQ ID NO: 265

            TPALLAPXEXNATXTANFPNIPVAXTXQPAXILGGXHXEGLLQIAXGPQPGQQQNGFTGQ SEQ ID NO: 263
            TPALLAP.E.NAT.TANFPNIPVA.T.QPA.ILGG.H.EGLLQIA.GPQPGQQQNGFTGQ SEQ ID NO: 264
            TPALLAPSESNATSTANFPNIPVASTSQPASILGGSHSEGLLQIASGPQPGQQQNGFTGQ SEQ ID NO: 265

            PATYHHNXTTTWTGXRTAPYTPNLPHHQKG     155 SEQ ID NO: 263
            PATYHHN.TTTWTG.RTAPYTPNLPHHQ.G         SEQ ID NO: 264
            PATYHHNSTTTWTGSRTAPYTPNLPHHQNG     286 SEQ ID NO: 265
```

Figure 42

Alignment: No86 - trembl|D32210|D32210_1

gene: "Notch2"; product: "cell surface protein"; Mus musculus (Notch2) mRNA, complete cds. //:gp|D32210|2373395 gene: "Notch2"; product: "cell surface protein"; Mus musculus (Notch2) mRNA, complete cds.

```
Q:      81 MPALRPALLWALLALWLCCATPAHALQCRDGYEPCVNEGMCVTYHNGTGYCKCP-GFLGE  SEQ ID NO: 266
           MP LRPA.L ALL LWLC A PAHALQCR.G.EPCVNEG.CVTYHNGTG:C:CP GFLGE  SEQ ID NO: 267
H:       1 MPDLRPAALRALLWLWLCGAGPAHALQCRGGQEPCVNEGTCVTYHNGTGFCRCPEGFLGE  SEQ ID NO: 268

          YCQHR-PCEKNRCGDPSTC     157  SEQ ID NO: 266
          YCQHR PCEKNRC : .TC          SEQ ID NO: 267
          YCQHRDPCEKNRCQNGGTC      79  SEQ ID NO: 268
```

Figure 43

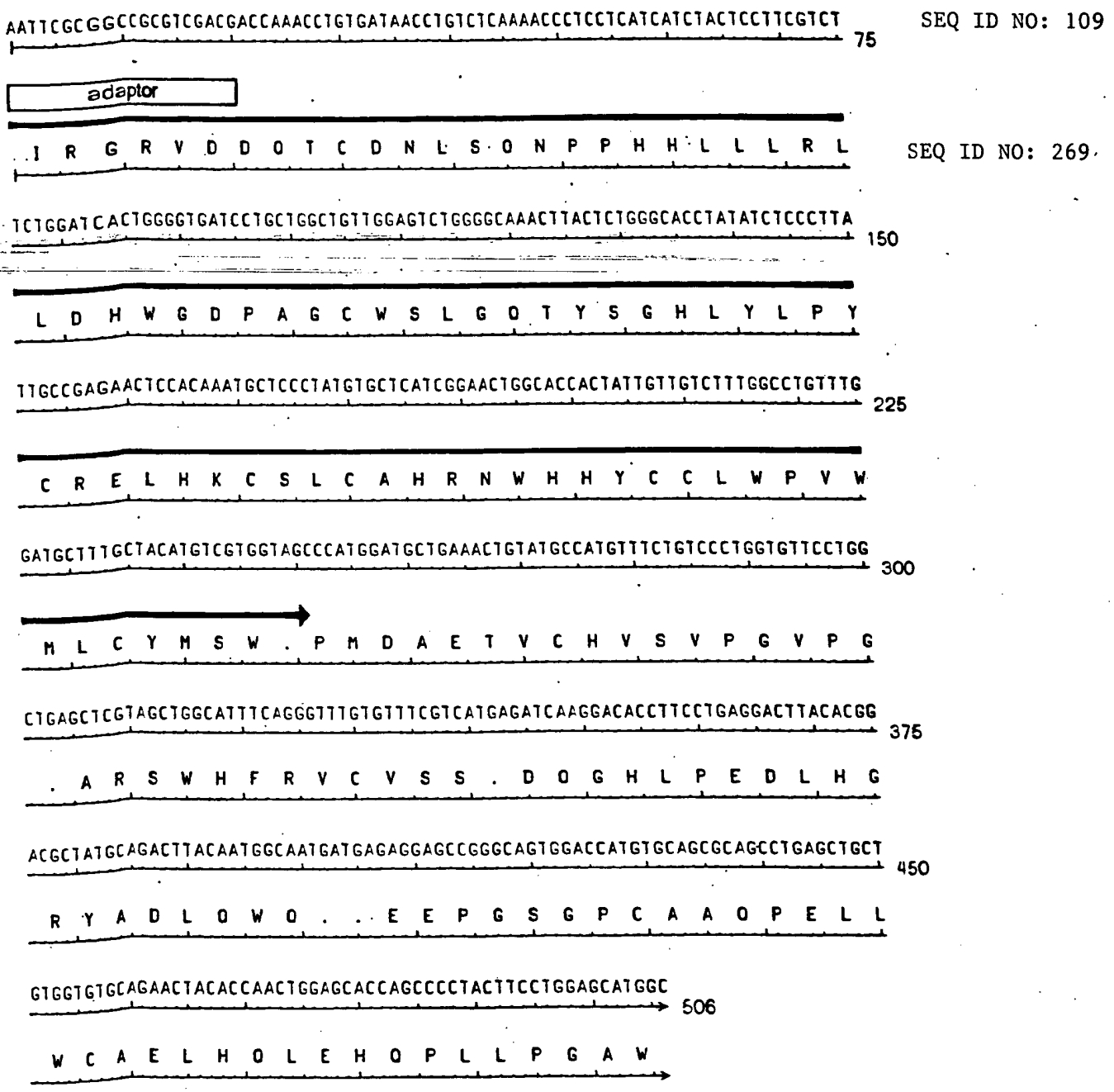
AATTCGCGGCCGCGTCGACGACCAAACCTGTGATAACCTGTCTCAAAACCCTCCTCATCATCTACTCCTTCGTCT
75
SEQ ID NO: 109
adaptor
I R G R V D D Q T C D N L S Q N P P H H L L L R L
SEQ ID NO: 269
TCTGGATCACTGGGGTGATCCTGCTGGCTGTTGGAGTCTGGGGCAAACTTACTCTGGGCACCTATATCTCCCTTA
150
L D H W G D P A G C W S L G Q T Y S G H L Y L P Y
TTGCCGAGAACTCCACAAATGCTCCCTATGTGCTCATCGGAACTGGCACCACTATTGTTGTCTTTGGCCTGTTTG
225
C R E L H K C S L C A H R N W H H Y C C L W P V W
GATGCTTTGCTACATGTCGTGGTAGCCCATGGATGCTGAAACTGTATGCCATGTTTCTGTCCCTGGTGTTCCTGG
300
M L C Y M S W . P M D A E T V C H V S V P G V P G
CTGAGCTCGTAGCTGGCATTTCAGGGTTTGTGTTTCGTCATGAGATCAAGGACACCTTCCTGAGGACTTACACGG
375
. A R S W H F R V C V S S . D Q G H L P E D L H G
ACGCTATGCAGACTTACAATGGCAATGATGAGAGGAGCCGGGCAGTGGACCATGTGCAGCGCAGCCTGAGCTGCT
450
R Y A D L Q W Q . . E E P G S G P C A A Q P E L L
GTGGTGTGCAGAACTACACCAACTGGAGCACCAGCCCCTACTTCCTGGAGCATGGC
506
W C A E L H Q L E H Q P L L P G A W

Figure 44

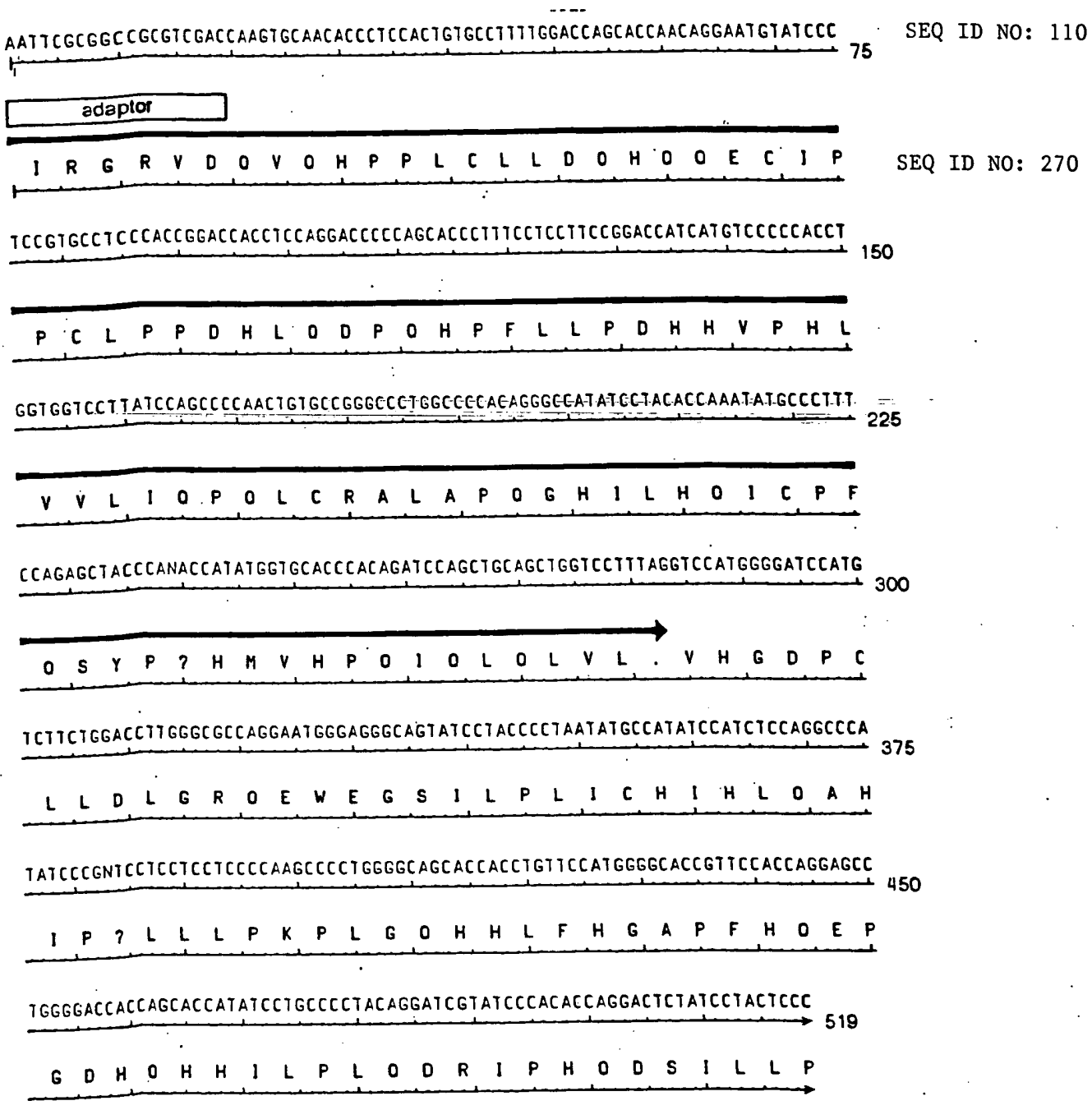

Figure 45

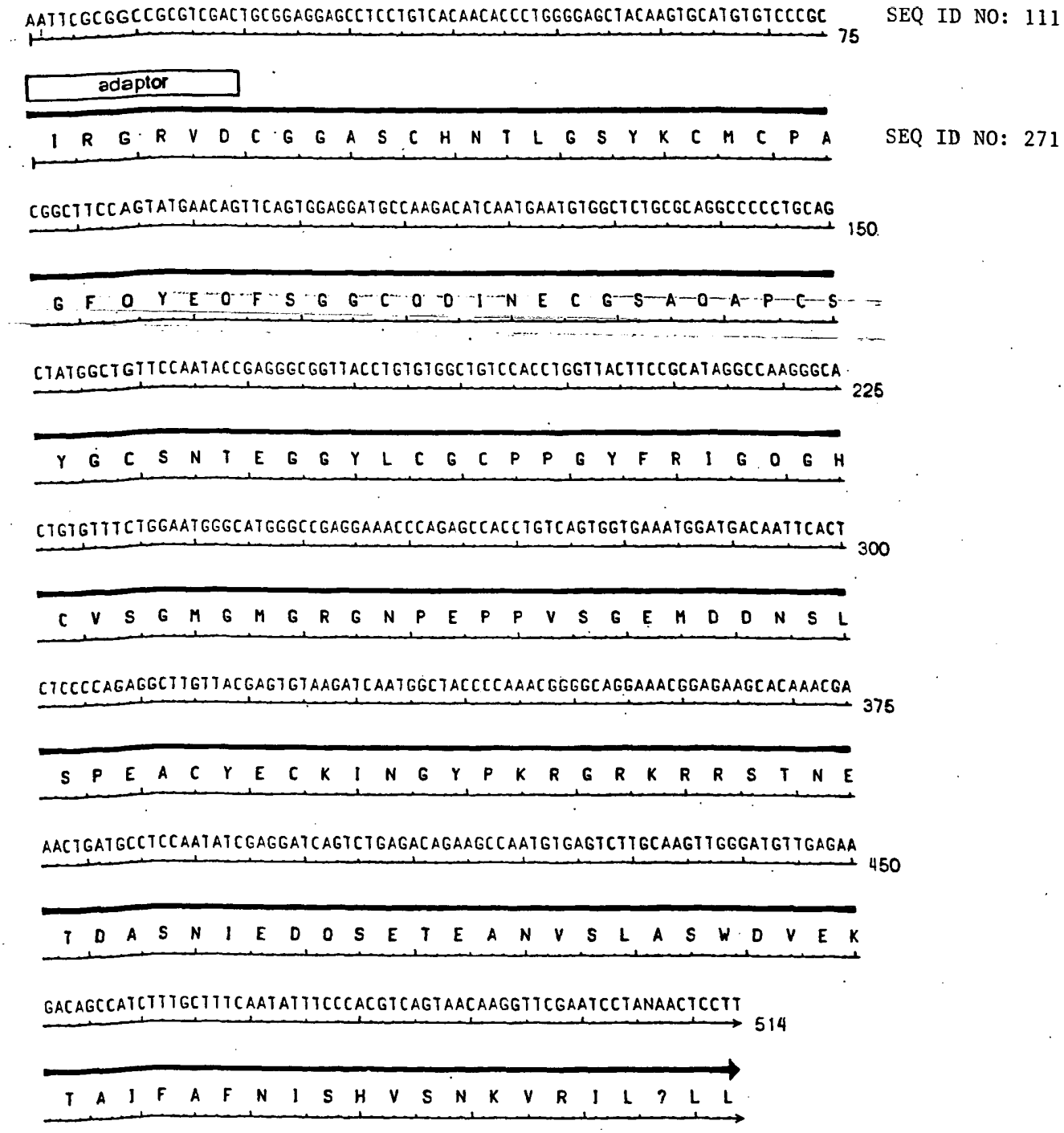
AATTCGCGGCCGCGTCGACTGCGGAGGAGCCTCCTGTCACAACACCCTGGGGAGCTACAAGTGCATGTGTCCCGC
SEQ ID NO: 111
75
adaptor
I R G R V D C G G A S C H N T L G S Y K C M C P A
SEQ ID NO: 271
CGGCTTCCAGTATGAACAGTTCAGTGGAGGATGCCAAGACATCAATGAATGTGGCTCTGCGCAGGCCCCCTGCAG
150
G F Q Y E Q F S G G C Q D I N E C G S A Q A P C S
CTATGGCTGTTCCAATACCGAGGGCGGTTACCTGTGTGGCTGTCCACCTGGTTACTTCCGCATAGGCCAAGGGCA
225
Y G C S N T E G G Y L C G C P P G Y F R I G Q G H
CTGTGTTTCTGGAATGGGCATGGGCCGAGGAAACCCAGAGCCACCTGTCAGTGGTGAAATGGATGACAATTCACT
300
C V S G M G M G R G N P E P P V S G E M D D N S L
CTCCCCAGAGGCTTGTTACGAGTGTAAGATCAATGGCTACCCCAAACGGGGCAGGAAACGGAGAAGCACAAACGA
375
S P E A C Y E C K I N G Y P K R G R K R R S T N E
AACTGATGCCTCCAATATCGAGGATCAGTCTGAGACAGAAGCCAATGTGAGTCTTGCAAGTTGGGATGTTGAGAA
450
T D A S N I E D Q S E T E A N V S L A S W D V E K
GACAGCCATCTTTGCTTTCAATATTTCCCACGTCAGTAACAAGGTTCGAATCCTANAACTCCTT
514
T A I F A F N I S H V S N K V R I L ? L L

Figure 46

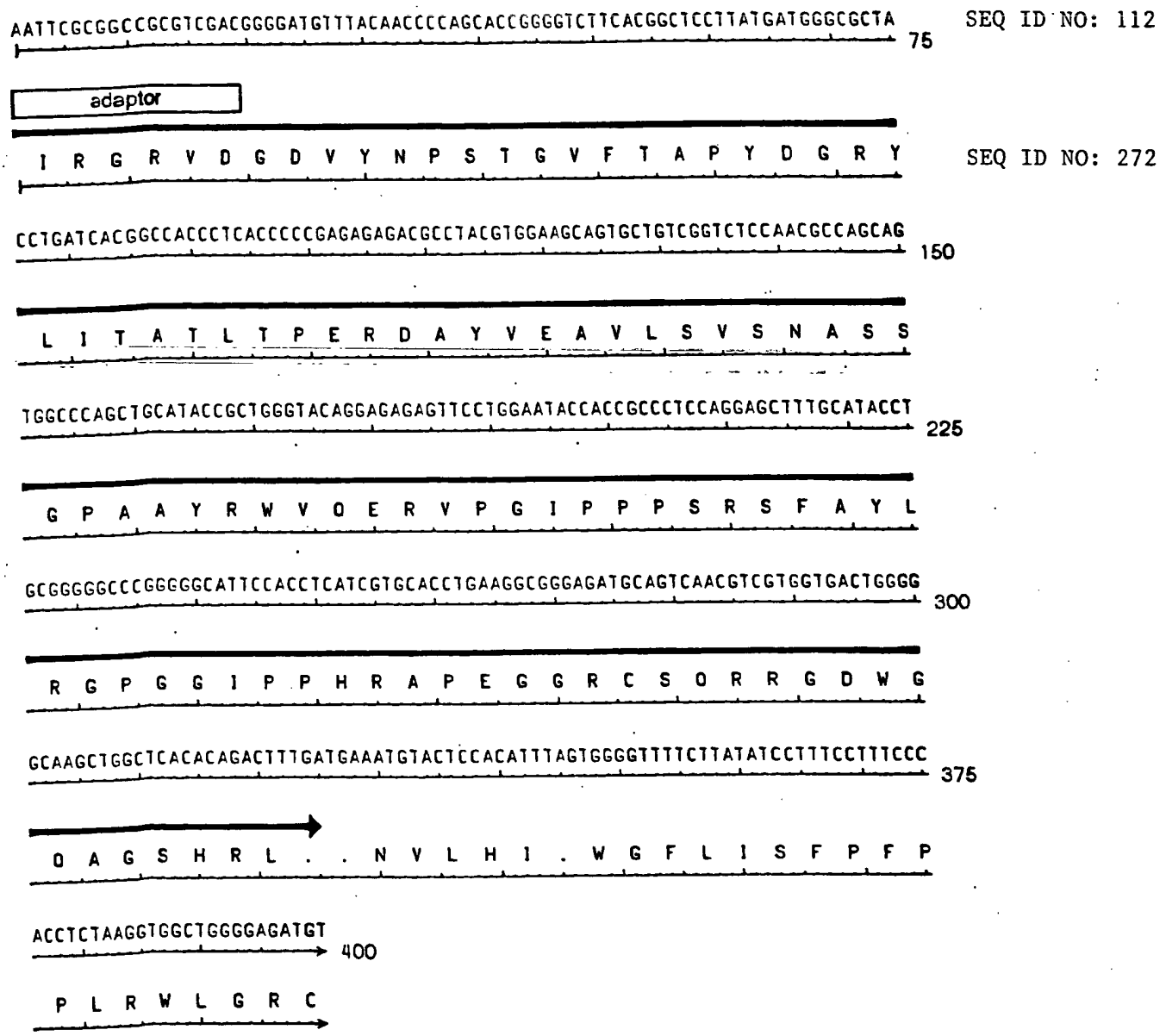
AATTCGCGGCCGCGTCGACGGGGATGTTTACAACCCCAGCACCGGGGTCTTCACGGCTCCTTATGATGGGCGCTA
SEQ ID NO: 112
75
adaptor
I R G R V D G D V Y N P S T G V F T A P Y D G R Y
SEQ ID NO: 272
CCTGATCACGGCCACCCTCACCCCCGAGAGAGACGCCTACGTGGAAGCAGTGCTGTCGGTCTCCAACGCCAGCAG
150
L I T A T L T P E R D A Y V E A V L S V S N A S S
TGGCCCAGCTGCATACCGCTGGGTACAGGAGAGAGTTCCTGGAATACCACCGCCCTCCAGGAGCTTTGCATACCT
225
G P A A Y R W V Q E R V P G I P P P S R S F A Y L
GCGGGGGCCCGGGGGCATTCCACCTCATCGTGCACCTGAAGGCGGGAGATGCAGTCAACGTCGTGGTGACTGGGG
300
R G P G G I P P H R A P E G G R C S Q R R G D W G
GCAAGCTGGCTCACACAGACTTTGATGAAATGTACTCCACATTTAGTGGGGTTTTCTTATATCCTTTCCTTTCCC
375
Q A G S H R L . . N V L H I . W G F L I S F P F P
ACCTCTAAGGTGGCTGGGGAGATGT
400
P L R W L G R C

Figure 47

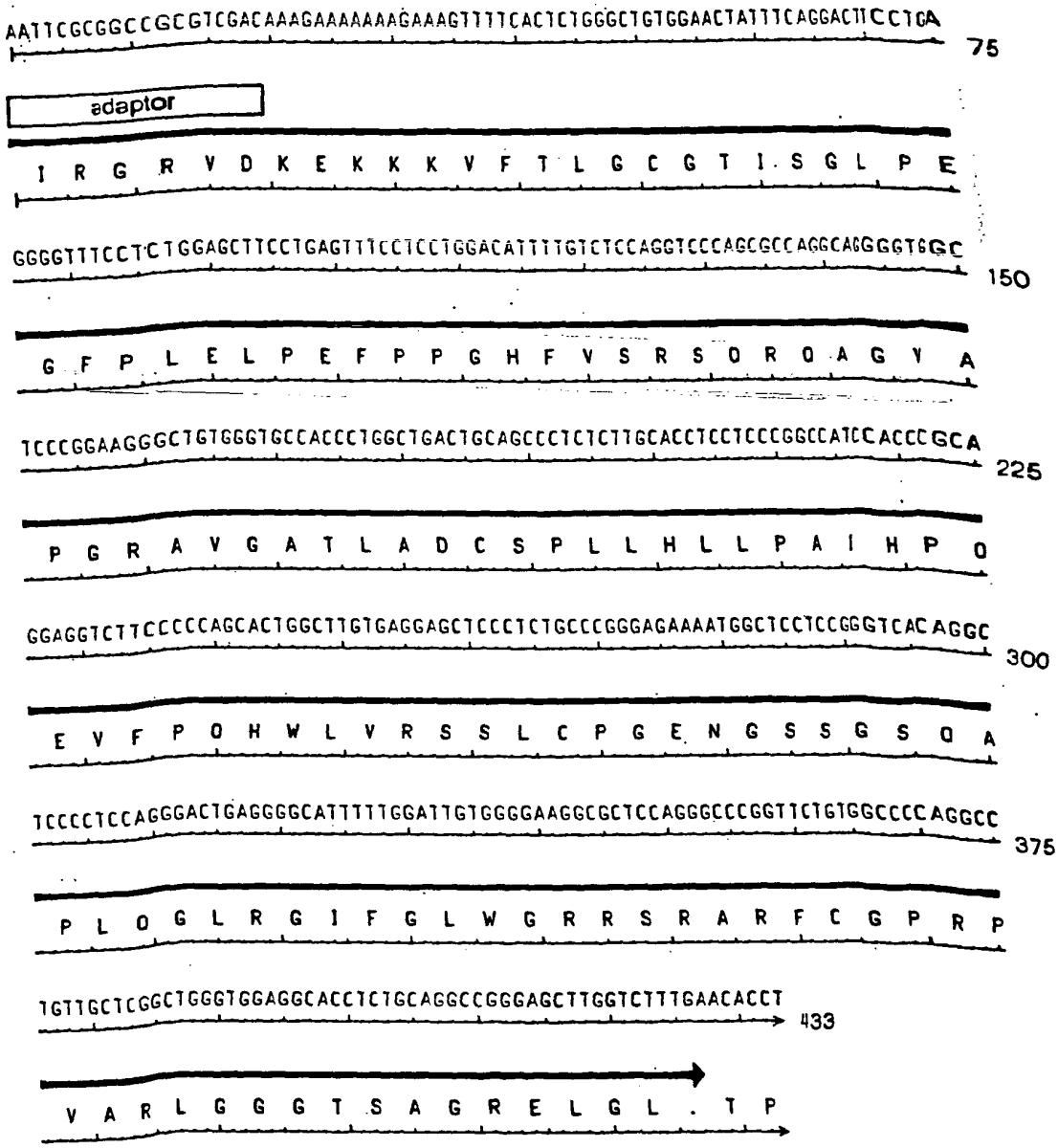
SEQ ID NO: 113
75
adaptor
I R G R V D K E K K K V F T L G C G T I S G L P E
SEQ ID NO: 273
150
G F P L E L P E F P P G H F V S R S Q R Q A G Y A
225
P G R A V G A T L A D C S P L L H L L P A I H P Q
300
E V F P Q H W L V R S S L C P G E N G S S G S Q A
375
P L Q G L R G I F G L W G R R S R A R F C G P R P
433
V A R L G G G T S A G R E L G L . T P

Figure 48

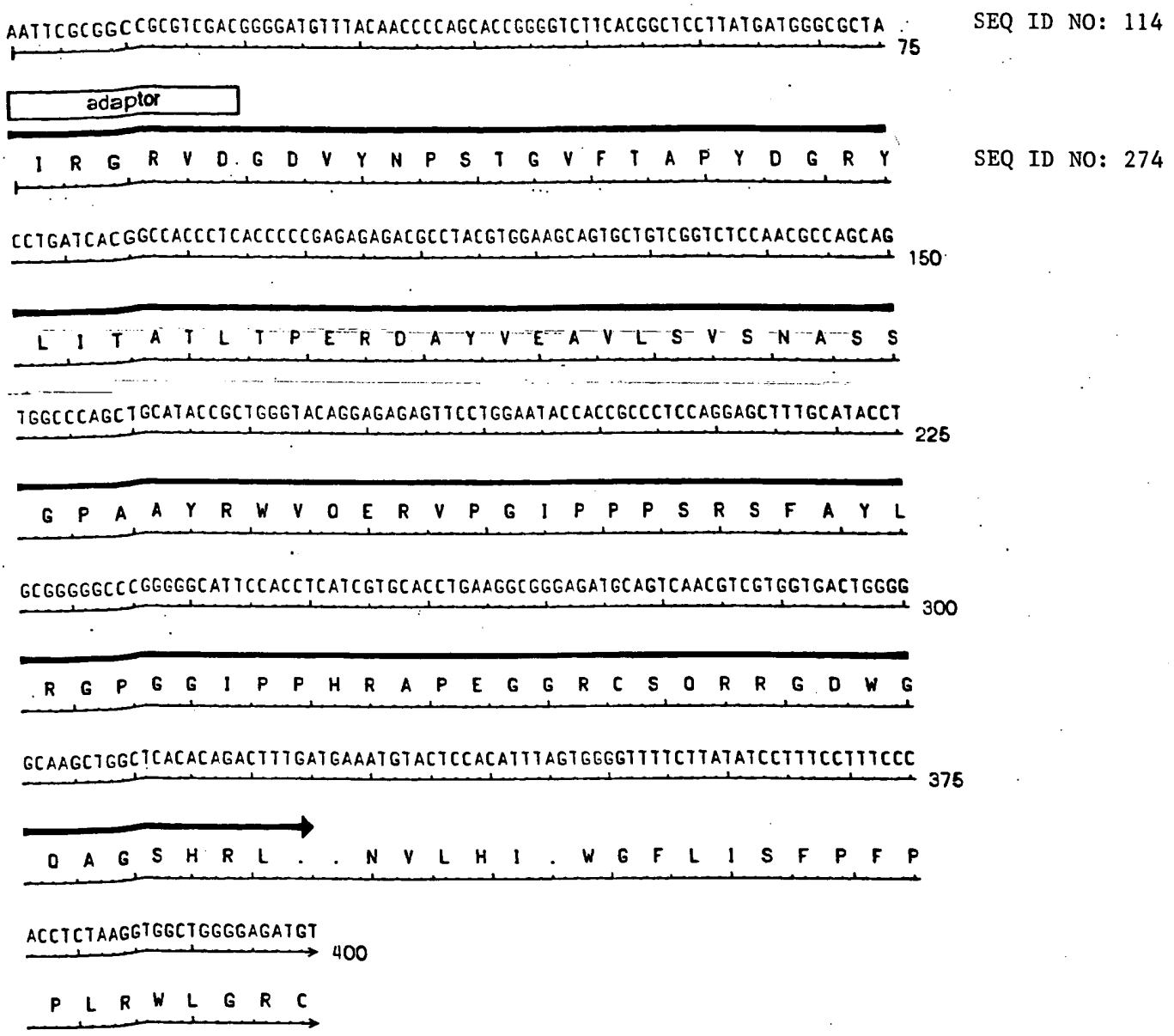
AATTCGCGGCCGCGTCGACGGGGATGTTTACAACCCCAGCACCGGGGTCTTCACGGCTCCTTATGATGGGCGCTA
75
SEQ ID NO: 114
adaptor
I R G R V D G D V Y N P S T G V F T A P Y D G R Y
SEQ ID NO: 274
CCTGATCACGGCCACCCTCACCCCCGAGAGAGACGCCTACGTGGAAGCAGTGCTGTCGGTCTCCAACGCCAGCAG
150
L I T A T L T P E R D A Y V E A V L S V S N A S S
TGGCCCAGCTGCATACCGCTGGGTACAGGAGAGAGTTCCTGGAATACCACCGCCCTCCAGGAGCTTTGCATACCT
225
G P A A Y R W V Q E R V P G I P P P S R S F A Y L
GCGGGGGCCCGGGGGCATTCCACCTCATCGTGCACCTGAAGGCGGGAGATGCAGTCAACGTCGTGGTGACTGGGG
300
R G P G G I P P H R A P E G G R C S Q R R G D W G
GCAAGCTGGCTCACACAGACTTTGATGAAATGTACTCCACATTTAGTGGGGTTTTCTTATATCCTTTCCTTTCCC
375
Q A G S H R L . . N V L H I . W G F L I S F P F P
ACCTCTAAGGTGGCTGGGGAGATGT
400
P L R W L G R C

Figure 49

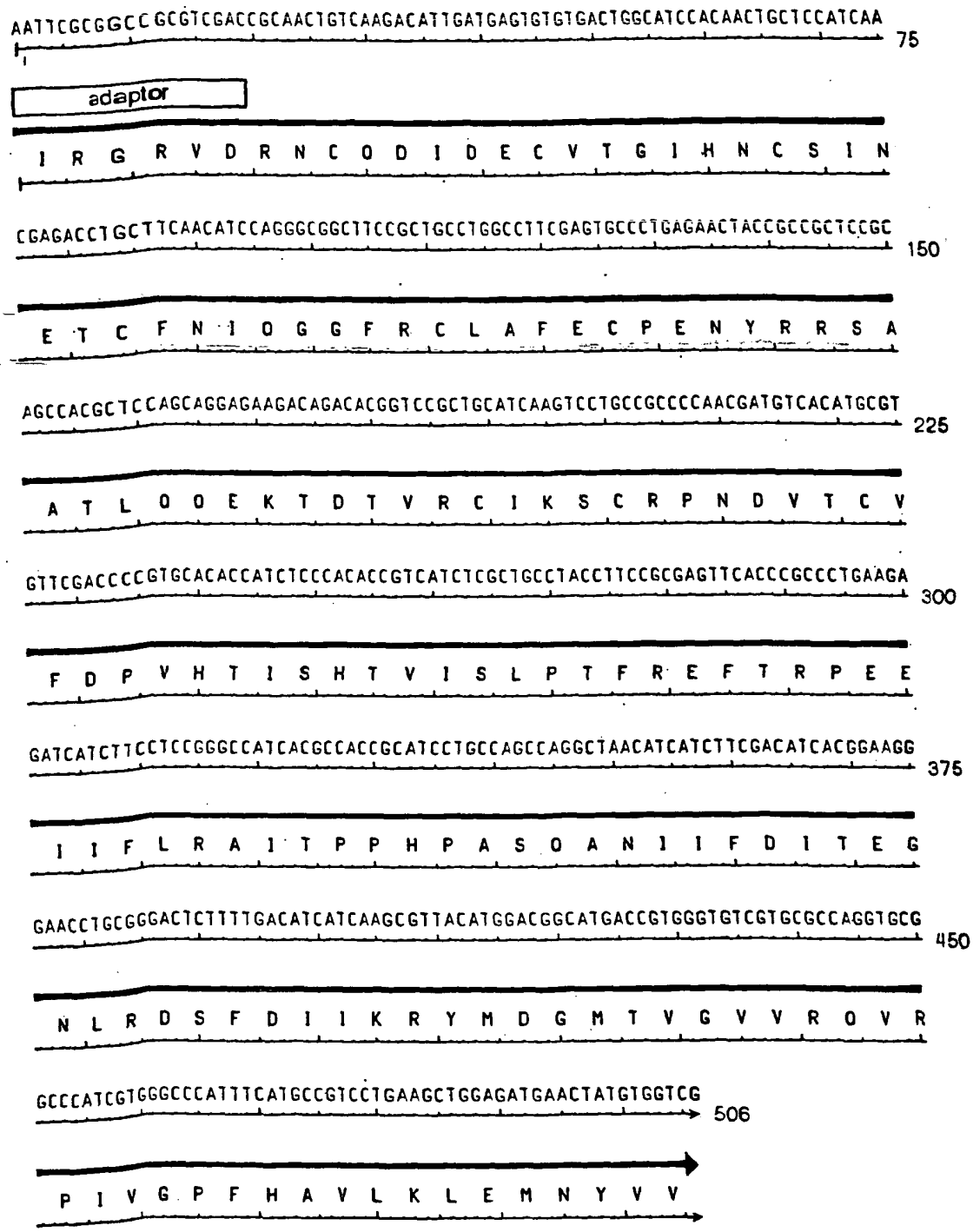
AATTCGCGGCCGCGTCGACCGCAACTGTCAAGACATTGATGAGTGTGTGACTGGCATCCACAACTGCTCCATCAA
75
SEQ ID NO: 115
adaptor
I R G R V D R N C Q D I D E C V T G I H N C S I N
SEQ ID NO: 275
CGAGACCTGCTTCAACATCCAGGGCGGCTTCCGCTGCCTGGCCTTCGAGTGCCCTGAGAACTACCGCCGCTCCGC
150
E T C F N I Q G G F R C L A F E C P E N Y R R S A
AGCCACGCTCCAGCAGGAGAAGACAGACACGGTCCGCTGCATCAAGTCCTGCCGCCCCAACGATGTCACATGCGT
225
A T L Q Q E K T D T V R C I K S C R P N D V T C V
GTTCGACCCCGTGCACACCATCTCCCACACCGTCATCTCGCTGCCTACCTTCCGCGAGTTCACCCGCCCTGAAGA
300
F D P V H T I S H T V I S L P T F R E F T R P E E
GATCATCTTCCTCCGGGCCATCACGCCACCGCATCCTGCCAGCCAGGCTAACATCATCTTCGACATCACGGAAGG
375
I I F L R A I T P P H P A S Q A N I I F D I T E G
GAACCTGCGGGACTCTTTTGACATCATCAAGCGTTACATGGACGGCATGACCGTGGGTGTCGTGCGCCAGGTGCG
450
N L R D S F D I I K R Y M D G M T V G V V R Q V R
GCCCATCGTGGGCCCATTTCATGCCGTCCTGAAGCTGGAGATGAACTATGTGGTCG
506
P I V G P F H A V L K L E M N Y V V

Figure 50

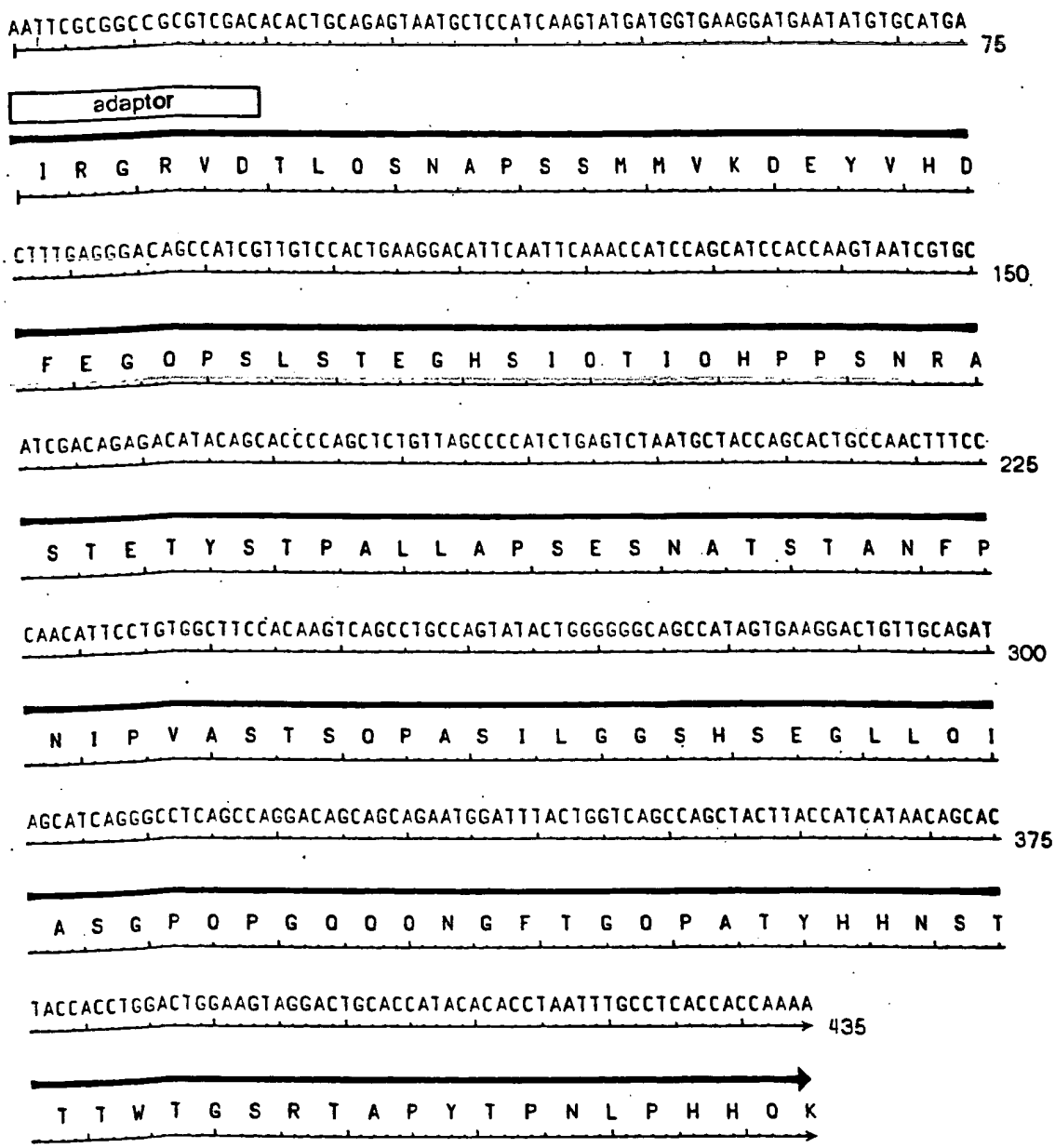
AATTCGCGGCCGCGTCGACACACTGCAGAGTAATGCTCCATCAAGTATGATGGTGAAGGATGAATATGTGCATGA
75
SEQ ID NO: 116
adaptor
I R G R V D T L Q S N A P S S M M V K D E Y V H D
SEQ ID NO: 276
CTTTGAGGGACAGCCATCGTTGTCCACTGAAGGACATTCAATTCAAACCATCCAGCATCCACCAAGTAATCGTGC
150
F E G Q P S L S T E G H S I Q T I Q H P P S N R A
ATCGACAGAGACATACAGCACCCCAGCTCTGTTAGCCCCATCTGAGTCTAATGCTACCAGCACTGCCAACTTTCC
225
S T E T Y S T P A L L A P S E S N A T S T A N F P
CAACATTCCTGTGGCTTCCACAAGTCAGCCTGCCAGTATACTGGGGGGCAGCCATAGTGAAGGACTGTTGCAGAT
300
N I P V A S T S Q P A S I L G G S H S E G L L Q I
AGCATCAGGGCCTCAGCCAGGACAGCAGCAGAATGGATTTACTGGTCAGCCAGCTACTTACCATCATAACAGCAC
375
A S G P Q P G Q Q Q N G F T G Q P A T Y H H N S T
TACCACCTGGACTGGAAGTAGGACTGCACCATACACACCTAATTTGCCTCACCACCAAAA
435
T T W T G S R T A P Y T P N L P H H Q K

Figure 51

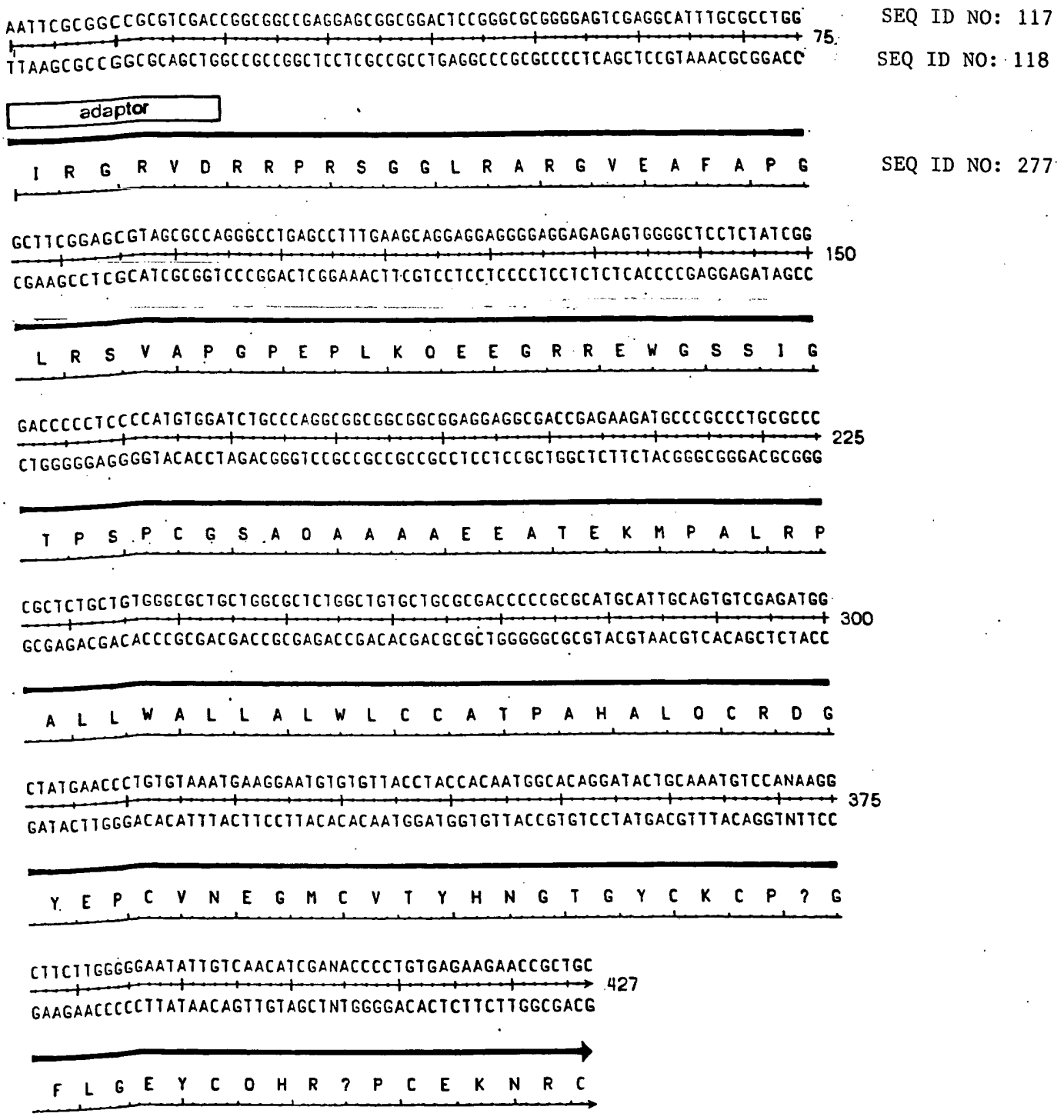
AATTCGCGGCCGCGTCGACCGGCGGCCGAGGAGCGGCGGACTCCGGGCGCGGGGAGTCGAGGCATTTGCGCCTGG
75
TTAAGCGCCGGCGCAGCTGGCCGCCGGCTCCTCGCCGCCTGAGGCCCGCGCCCCTCAGCTCCGTAAACGCGGACC
SEQ ID NO: 117
SEQ ID NO: 118
adaptor
I R G R V D R R P R S G G L R A R G V E A F A P G
SEQ ID NO: 277
GCTTCGGAGCGTAGCGCCAGGGCCTGAGCCTTTGAAGCAGGAGGAGGGGAGGAGAGAGTGGGGCTCCTCTATCGG
150
CGAAGCCTCGCATCGCGGTCCCGGACTCGGAAACTTCGTCCTCCTCCCCTCCTCTCTCACCCCGAGGAGATAGCC
L R S V A P G P E P L K Q E E G R R E W G S S I G
GACCCCCTCCCCATGTGGATCTGCCCAGGCGGCGGCGGCGGAGGAGGCGACCGAGAAGATGCCCGCCCTGCGCCC
225
CTGGGGGAGGGGTACACCTAGACGGGTCCGCCGCCGCCGCCTCCTCCGCTGGCTCTTCTACGGGCGGGACGCGGG
T P S P C G S A Q A A A A E E A T E K M P A L R P
CGCTCTGCTGTGGGCGCTGCTGGCGCTCTGGCTGTGCTGCGCGACCCCCGCGCATGCATTGCAGTGTCGAGATGG
300
GCGAGACGACACCCGCGACGACCGCGAGACCGACACGACGCGCTGGGGGCGCGTACGTAACGTCACAGCTCTACC
A L L W A L L A L W L C C A T P A H A L Q C R D G
CTATGAACCCTGTGTAAATGAAGGAATGTGTGTTACCTACCACAATGGCACAGGATACTGCAAATGTCCANAAGG
375
GATACTTGGGACACATTTACTTCCTTACACACAATGGATGGTGTTACCGTGTCCTATGACGTTTACAGGTNTTCC
Y E P C V N E G M C V T Y H N G T G Y C K C P ? G
CTTCTTGGGGGAATATTGTCAACATCGANACCCCTGTGAGAAGAACCGCTGC
427
GAAGAACCCCCTTATAACAGTTGTAGCTNTGGGGACACTCTTCTTGGCGACG
F L G E Y C Q H R ? P C E K N R C

Figure 52

Pathogenicity Island Probe # (bp #)	*P. aeruginosa* srains that gave Positive signal with the specific probe used	*P. aeruginosa* srains that gave Negative signal with the specific probe used
3 (25562-26456)	PA14, CF2, CF6, CF26, CF29	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF30, CF32
4 (61181-63607)	PA14, CF2, CF6, CF26, CF29	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF30, CF32
5 (74933-76117)	PA14, PAO37, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
6 (84922-86622)	PAO1, PA14, PAO37, CF2, CF, CF26,	PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
7 (103070-104556)	PA14, PAO37, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
8 (104799-105545)	PA14, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32

These experiments indicate that at least part of the big island region contained in each probe is present in the *P. aeruginosa* strains that gave positive hybridization signal.

Figure 53

ORF7 Protein SEQ ID NO: 278

MINSHLLYRLSYRGTSFFQPWTLPVLLDSRLRGAPFYGCARACQPSDPKSFSSFSTSDKTALPLHAAALSRLPDAHEKAP
PKRGFPCPPPKRSGEDDLVAFHLRRDTGTRREFAGQDQLRQRVLDPALDGPLQRACAIDRVEADGNQLVQRLLAQFQAQL
ALGQALAQATELDLGDAGDLLASQRLEHHHFVDPVDEFRTEVRIDRVHHCGTLRLAVAGQLLDLRRTEVGGHHHHGVAEV
HRTPVTVGQASVLEHLEENVEYIRMGLLHLVQQHHRVGLAADRLGQVAAFLEADVARRRADQAGHRVFLHELGHIYPHQR
LLGIEEELGQRLAQLGLAHPGRAEEEERAARPVRIGEAGARTAHGVGHGDYRLVLADHSPMQLLLHAQQLLALALEHLRH
RDTGPLGNHFGDFLVGHLVAQQLVLGLAVLVDHLQAAFQVRDGLVLDARHALEVALAPRRLHLLLGLLDLLLDLRRALHL
GLLGLPDLLEVGVFALELDDILLQLGQALPGGFVVFLLQRLALDLQLDQATVETIQFLRLGVDLHADAAGGLVDQVDGLV
RQLPIGDVAVRQLGRGDDRAVGDAHPVVHFIAFLEATEDGDGVFLARFVHQHLLEAALQRGILLDVLAILVEGSSTDAVQ
LAARQSRLEHVAGVHGTFRLAGADHGVQFVDEQDDPAFLLAQFVEDRLQAFLELAAELGTGDQRPHVQGQQALVLEAVRH
FAVDDALGQALDDGGLADAGFADQHRVVLGPPLQDLDGPADLVVATDHRVELAFLGALGHVDGVLVQRLARLLDVRVVHR
FAATQVGHGILQRLARHALAEQQLAEPGVLVHRGQQYQLAGDELVALLLGQAVSLVEQACEILGQVHVAGRALDLRQRVE
FFVEAAAQGGDIEADLHQQGLDRTALLLEQGGKQVHRLDGRMVMANGQGLGVGERQLQLAGQTVYSHGSSFLL.

Figure 54

ORF7 nucleotide sequence SEQ ID NO: 119

ATGATTAACAGTCATTTGCTCTACCGACTGAGCTATCGCGGAACGTCTTTCTTCCAACCCTGGACGCTTCCGGTGTTGCT
GGATTCGCGTCTCAGAGGCGCGCCATTTTACGGATGCGCGCGGGCATGTCAACCCTCTGATCCAAAAAGTTTTTCTTCTT
TTTCCACGAGCGACAAAACGGCCCTTCCACTGCATGCGGCAGCGCTCTCGCGCCTACCGGACGCCCATGAAAAAGCCCCG
CCGAAGCGGGGCTTTCCCTGTCCGCCCCCGAAGAGGTCAGGCGAAGACGATCTCGTCGCCTTCCACCTTCGCCGAGATAC
TGGCACCCGGCGCGAATTTGCCGGCCAGGATCAGTTGCGCCAGCGGGTTCTCGATCCAGCGCTGGATGGCCCGCTTCAGC
GGGCGTGCGCCATAGACCGGGTCGAAGCCGACGGCAATCAGCTTGTCCAGCGCCTCCTGGCTCAGTTCCAGGCTCAGCTC
GCGCTCGGCCAGGCGCTTGCGCAGGCGACCGAGCTGGATCTCGGCGATGCCGGCGATCTGCTCGCGAGCCAGCGGCTCGA
ACACCACCACTTCGTCGATCCGGTTGATGAATTCCGGACGGAAGTGCGCATTGACCGCGTCCATCACTGCGGCACGTTGC
GCCTCGCGGTCGCCGGCCAGCTCCTGGATCTGCGCCGAACCGAGGTTGGAGGTCATCACCACCACGGTGTTGCGGAAGTC
CACCGTACGCCCGTGACTGTCGGTCAGGCGTCCGTCCTCGAGCACCTGGAGGAGAATGTTGAATACATCCGGATGGGCCT
TCTCCACCTCGTCCAGCAGCACCACCGAGTAGGGCTTGCGGCGGATCGCCTCGGTCAGGTAGCCGCCTTCCTCGAAGCCG
ACGTAGCCCGGAGGCGCGCCGATCAGGCGGGCCACCGAGTGTTTCTCCATGAACTCGGACATATCTATCCGCACCAGCGC
CTCCTCGGTATCGAAGAGGAACTCGGCCAG
CGCCTTGCACAACTCGGTCTTGCCCACCCCGGTCGGGCCGAGGAAGAGGAACGAGCCGCTCGGCCGGTTCGGATCGGCGA
GGCCGGCGCGCGAACGGCGCACGGCGTTGGACACGGCGACTACCGCCTCGTCCTGGCCGATCACTCGCCGATGCAGCTCC
TGCTCCATGCGCAGCAGCTTCTCGCGCTCGCCCTCGAGCATCTTCGACACCGGGATACCGGTCCACTTGGAAACCACTTC
GGCGATTTCCTCGTCGGTCACCTTGTTGCGCAGCAACTGGTTCTCGGTCTTGCCGTGCTGGTCGACCATCTGCAGGCTGC
GTTCCAGGTCCGGGATGGTCTGGTACTGGATGCGCGCCATGCTCTCGAGGTCGCCCTTGCGCCGCGCCGCCTCCATCTCC
TGCTTGGCCTGCTCGATCTTCTGCTGGATCTGCGCCGAGCCCTGCACCTCGGCCTTCTCGGACTTCCAGATCTCCTCGAG
GTCGGCGTATTCGCGCTCGAGCTTGACGATATCCTCCTCCAGCTTGGCCAGGCGCTTCCTGGTGGCTTCGTCGTCTTCCT
TCTTCAGCGCCTCGCGCTCGATCTTCAGCTGGATCAGGCGACGGTCGAGACGATCCAGTTCCTCCGGCTTGGAGTCGATC
TCCATGCGGATGCGGCTGGCGGCCTCGTCGATCAGGTCGATGGCCTTGTCCGGCAGTTGCCGATCGGTGATGTAGCGGTG
CGACAGCTTGGCCGCGGCGATGATCGCGCCGTCGGTGATGCTCACCCCGTGGTGCACTTCATAGCGTTCCTTGAGGCCAC
GGAGGATGGCGATGGTGTCTTCCTCGCTCGGTTCGTCCACCAGCACCTTCTGGAAGCGGCGCTCCAGCGCGGCATCCTTC
TCGATGTACTGGCGATACTCGTCGAGGGTAGTAGCACCGACGCAGTGCAGCTCGCCGCGCGCCAGAGCCGGCTTGAGCAT
GTTGCCGGCGTCCATGGCACCTTCCGCCTT
GCCGGCGCCGACCATGGTGTGCAGTTCGTCGATGAACAGGATGACCCGGCCTTCCTGCTTGCCCAGTTCGTTGAGGACCG
CCTTCAGGCGTTCCTCGAACTCGCCGCGGAACTTGGCACCGGCGATCAGCGCCCCCATGTCCAGGGCCAGCAGGCGCTTG
TCCTTGAGGCCGTCCGGCACTTCGCCGTTGATGATGCGCTGGGCCAGGCCCTCGACGATGGCGGTCTTGCCGACGCCGGG
TTCGCCGATCAGCACCGGGTTGTTCTTGGTCCGCCGCTGCAGGACCTGGATGGTCCGGCGGATCTCGTCGTCGCGACCGA
TCACCGGGTCGAGCTTGCCTTCCTCGGCGCGCTTGGTCATGTCGACGGTGTACTTGTCCAGCGCCTGGCGCGACTCCTCG
ACGTTCGGGTCGTTCACCGCTTCGCCGCCACGCAGGTTGGCCACGGCATTCTCCAGCGCCTTGCGCGACACGCCCTGGCC
GAGCAGCAGCTTGCCGAGCCTGGTGTTCTCGTCCATCGCGGCCAGCAATACCAGCTCGCTGGAGATGAACTGGTCGCCCT
TCTGCTGGGCCAGGCGGTCAGCCTGGTTGAGCAGGCGTGCGAGATCCTGGGACAGGTTCACGTCGCCGGTCGGGCTCTGG
ATCTTCGGCAGCGCGTCGAGTTCTTTGTTGAGGCCGCTGCGCAGGGCGGCGATATCGAAGCCGACCTGCATCAGCAGGGG
CTTGATCGAACCGCCTTGCTGCTCGAGCAGGGCGGAAAGCAGGTGCACCGGCTCGATGGCCGGATGGTCATGGCCAACGG
CCAGGGACTGGGCGTCGGAGAGCGCCAGTTGCAGCTTGCTGGTCAAACGGTCTATTCGCATGGGTCGTCCTTCCTTCTAT
AG

Figure 55

clpB protein SEQ ID NO: 279

MRIDRLTSKLQLALSDAQSLAVGHDHPAIEPVHLLSALLEQQGGSIKPLLMQVGFDIAALRSGLNKELDALPKIQSPTGD
VNLSQDLARLLNQADRLAQQKGDQFISSELVLLAAMDENTRLGKLLLGQGVSRKALENAVANLRGGEAVNDPNVEESRQA
LDKYTVDMTKRAEEGKLDPVIGRDDEIRRTIQVLQRRTKNNPVLIGEPGVGKTAIVEGLAQRIINGEVPDGLKDKRLLAL
DMGALIAGAKFRGEFEERLKAVLNELGKQEGRVILFIDELHTMVGAGKAEGAMDAGNMLKPALARGELHCVGATTLDEYR
QYIEKDAALERRFQKVLVDEPSEEDTIAILRGLKERYEVHHGVSITDGAIIAAAKLSHRYITDRQLPDKAIDLIDEAASR
IRMEIDSKPEELDRLDRRLIQLKIEREALKKEDDEATRKRLAKLEEDIVKLEREYADLEEIWKSEKAEVQGSAQIQQKIE
QAKQEMEAARRKGDLESMARIQYQTIPDLERSLQMVDQHGKTENQLLRNKVTDEEIAEVVSKWTGIPVSKMLEGEREKLL
RMEQELHRRVIGQDEAVVAVSNAVRRSRAGLADPNRPSGSFLFLGPTGVGKTELCKALAEFLFDTEEALVRIDMSEFMEK
HSVARLIGAPPGYVGFEEGGYLTEAIRRKPYSVVLLDEVEKAHPDVFNILLQVLEDGRLTDSHGRTVDFRNTVVVMTSNL
GSAQIQELAGDREAQRAAVMDAVNAHFRPEFINRIDEVVVFEPLAREQIAGIAEIQLGRLRKRLAERELSLELSQEALDK
LIAVGFDPVYGARPLKRAIQRWIENPLAQLILAGKFAPGASISAKVEGDEIVFA.

Figure 56

clpB DNA SEQ ID NO: 120

ATGCGAATAGACCGTTTGACCAGCAAGCTGCAACTGGCGCTCTCCGACGCCCAGTCCCTGGCCGTTGGCCATGACCATCC
GGCCATCGAGCCGGTGCACCTGCTTTCCGCCCTGCTCGAGCAGCAAGGCGGTTCGATCAAGCCCCTGCTGATGCAGGTCG
GCTTCGATATCGCCGCCCTGCGCAGCGGCCTCAACAAAGAACTCGACGCGCTGCCGAAGATCCAGAGCCCGACCGGCGAC
GTGAACCTGTCCCAGGATCTCGCACGCCTGCTCAACCAGGCTGACCGCCTGGCCCAGCAGAAGGGCGACCAGTTCATCTC
CAGCGAGCTGGTATTGCTGGCCGCGATGGACGAGAACACCAGGCTCGGCAAGCTGCTGCTCGGCCAGGGCGTGTCGCGCA
AGGCGCTGGAGAATGCCGTGGCCAACCTGCGTGGCGGCGAAGCGGTGAACGACCCGAACGTCGAGGAGTCGCGCCAGGCG
CTGGACAAGTACACCGTCGACATGACCAAGCGCGCCGAGGAAGGCAAGCTCGACCCGGTGATCGGTCGCGACGACGAGAT
CCGCCGGACCATCCAGGTCCTGCAGCGGCGGACCAAGAACAACCCGGTGCTGATCGGCGAACCCGGCGTCGGCAAGACCG
CCATCGTCGAGGGCCTGGCCCAGCGCATCATCAACGGCGAAGTGCCGGACGGCCTCAAGGACAAGCGCCTGCTGGCCCTG
GACATGGGGGCGCTGATCGCCGGTGCCAAGTTCCGCGGCGAGTTCGAGGAAGGCCTGAAGGCGGTCCTCAACGAACTGGG
CAAGCAGGAAGGCCGGGTCATCCTGTTCATCGACGAACTGCACACCATGGTCGGCGCCGGCAAGGCGGAAGGTGCCATGG
ACGCCGGCAACATGCTCAAGCCGGCTCTGGCGCGCGGCGAGCTGCACTGCGTCGGTGCTACTACCCTCGACGAGTATCGC
CAGTACATCGAGAAGGATGCCGCGCTGGAG
CGCCGCTTCCAGAAGGTGCTGGTGGACGAACCGAGCGAGGAAGACACCATCGCCATCCTCCGTGGCCTCAAGGAACGCTA
TGAAGTGCACCACGGGGTGAGCATCACCGACGGCGCGATCATCGCCGCGGCCAAGCTGTCGCACCGCTACATCACCGATC
GGCAACTGCCGGACAAGGCCATCGACCTGATCGACGAGGCCGCCAGCCGCATCCGCATGGAGATCGACTCCAAGCCGGAG
GAACTGGATCGTCTCGACCGTCGCCTGATCCAGCTGAAGATCGAGCGCGAGGCGCTGAAGAAGGAAGACGACGAAGCCAC
CAGGAAGCGCCTGGCCAAGCTGGAGGAGGATATCGTCAAGCTCGAGCGCGAATACGCCGACCTCGAGGAGATCTGGAAGT
CCGAGAAGGCCGAGGTGCAGGGCTCGGCGCAGATCCAGCAGAAGATCGAGCAGGCCAAGCAGGAGATGGAGGCGGCGCGG
CGCAAGGGCGACCTCGAGAGCATGGCGCGCATCCAGTACCAGACCATCCCGGACCTGGAACGCAGCCTGCAGATGGTCGA
CCAGCACGGCAAGACCGAGAACCAGTTGCTGCGCAACAAGGTGACCGACGAGGAAATCGCCGAAGTGGTTTCCAAGTGGA
CCGGTATCCCGGTGTCGAAGATGCTCGAGGGCGAGCGCGAGAAGCTGCTGCGCATGGAGCAGGAGCTGCATCGGCGAGTG
ATCGGCCAGGACGAGGCGGTAGTCGCCGTGTCCAACGCCGTGCGCCGTTCGCGCGCCGGCCTCGCCGATCCGAACCGGCC
GAGCGGCTCGTTCCTCTTCCTCGGCCCGACCGGGGTGGGCAAGACCGAGTTGTGCAAGGCGCTGGCCGAGTTCCTCTTCG
ATACCGAGGAGGCGCTGGTGCGGATAGATATGTCCGAGTTCATGGAGAAACACTCGGTGGCCCGCCTGATCGGCGCGCCT
CCGGGCTACGTCGGCTTCGAGGAAGGCGGC
TACCTGACCGAGGCGATCCGCCGCAAGCCCTACTCGGTGGTGCTGCTGGACGAGGTGGAGAAGGCCCATCCGGATGTATT
CAACATTCTCCTCCAGGTGCTCGAGGACGGACGCCTGACCGACAGTCACGGGCGTACGGTGGACTTCCGCAACACCGTGG
TGGTGATGACCTCCAACCTCGGTTCGGCGCAGATCCAGGAGCTGGCCGGCGACCGCGAGGCGCAACGTGCCGCAGTGATG
GACGCGGTCAATGCGCACTTCCGTCCGGAATTCATCAACCGGATCGACGAAGTGGTGGTGTTCGAGCCGCTGGCTCGCGA
GCAGATCGCCGGCATCGCCGAGATCCAGCTCGGTCGCCTGCGCAAGCGCCTGGCCGAGCGCGAGCTGAGCCTGGAACTGA
GCCAGGAGGCGCTGGACAAGCTGATTGCCGTCGGCTTCGACCCGGTCTATGGCGCACGCCCGCTGAAGCGGGCCATCCAG
CGCTGGATCGAGAACCCGCTGGCGCAACTGATCCTGGCCGGCAAATTCGCGCCGGGTGCCAGTATCTCGGCGAAGGTGGA
AGGCGACGAGATCGTCTTCGCCTGA

Figure 57

ORF ID	Strand	Left end	Right end	ORF length (aa)	G+C content (%)	Location prediction	Gene name	Gene function	Protein with the highest identity (Gene Name / Species Strain)	E-value (% identity)	GenBank accession no.	Proteins with lesser identity Cut off 30%
[illegible]	[illegible]	801	876	[illegible]	51	[illegible]	[illegible]	[illegible]	tRNA-lys / *P. aeruginosa* PAO1	[illegible]	AE004531	[illegible]
RS01	-	959	1,280	pseudogene	57			Hypothetical protein	PA0977 / *P. aeruginosa* PAO1	8E-37 (86)	AAG04366	
RS02	[illegible]	1,247	2,527	426	60	cytoplasm	*xerC*	Integrase	XerC / *P. aeruginosa* SG17M	0.0 (94)	AAG02084	STY4666
RS03	-	2,524	2,901	125	56	cytoplasm		Hypothetical protein	XF1753 / *X. fastidiosa* 9a5c	3E-34 (49)	AAF84562	XAC2196, XCC3121, STY4665
RS04	[illegible]	2,988	3,880	280	62	cytoplasm	[illegible]	Putative transposase	PA0978 / IS222 / *P. aeruginosa* PAO1	1E-154 (96)	AAG04367	PA1938, RSC2313
RS05	-	3,830	4,138	103	60	cytoplasm		Putative transposase	PA0979 / IS222 / *P. aeruginosa* PAO1	6E-50 (99)	AAG04368	PA1937, RSC2314
RS06	[illegible]	4,370	4,654	94	45	inner or outer membranes	[illegible]	Hypothetical protein	PA0980 / *P. aeruginosa* PAO1	4E-50 (97)	AAG04369	[illegible]
RS07	+	4,691	5,314	207	46	outer membrane and periplasm		Hypothetical protein	PA0981 / *P. aeruginosa* PAO1	E-114 (99)	AAG04370	
RS08	[illegible]	5,420	5,737	105	50	outer membrane and periplasm	[illegible]	Hypothetical protein	PA0673 / *P. aeruginosa* PAO1	2E-05 (37)	AAG04062	[illegible]
RS09	+	5,849	6,052	67	57	cytoplasm			No significant similarity			
RS10	[illegible]	6,108	6,326	72	58	cytoplasm	[illegible]	[illegible]	No significant similarity	[illegible]	[illegible]	[illegible]
RS11	+	6,540	6,989	149	53	inner membrance		Acetyltransferase	PP0651 / *P. putida* KT2440	2E-75 (95)	AAN66276	
RS12	[illegible]	7,118	8,213	366	59	outer membrane and periplasm	[illegible]	Transposase	PP3964 / ISPpu14 ORF3 / *P. putida* KT2440	0.0 (98)	AAN69558	[illegible]
RS13	+	8,253	8,822	189	61	cytoplasm		Putative transposase	PA0987 / *P. aeruginosa* PAO1	4E-88 (85)	AAG04376	XAC2424
RS14	[illegible]	8,959	11,022	687	59	cytoplasm	*exoU (pep4)*	cytotoxin (type III secretion system effector)	ExoU / *P. aeruginosa* PA103	0.0 (100)	AAC38269	[illegible]
RS15	+	11,019	11,432	137	56	inner membrance	*spcU*	ExoU chaperone	SpcU / *P. aeruginosa* PA103	4E-62 (94)	AAC16024	

FIGURE 58

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND PROTEINS AND USES THEREOF
Applicants: Laurence Rahme et al.
Filing Date: September 12, 2003 Serial No.: Not Yet Assigned
Page 105 of 118 Customer No.: 21559

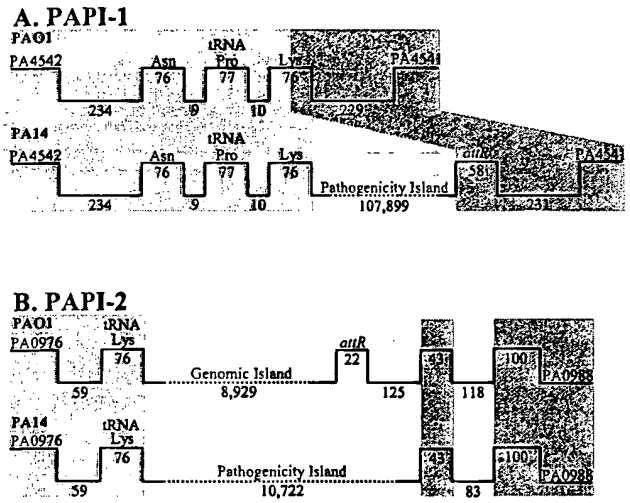
A. PAPI-1
PAO1
PA4542
Asn
tRNA
Pro
Lys
PA4541
76
77
76
234
9
10
229
PA14
PA4542
Asn
tRNA
Pro
Lys
76
77
76
Pathogenicity Island
107,899
234
9
10
58
231
PA4541
B. PAPI-2
PAO1
PA0976
tRNA
Lys
76
59
Genomic Island
8,929
attR
22
125
43
118
100
PA0988
PA14
PA0976
tRNA
Lys
76
59
Pathogenicity Island
10,722
43
83
100
PA0988

Figure 59

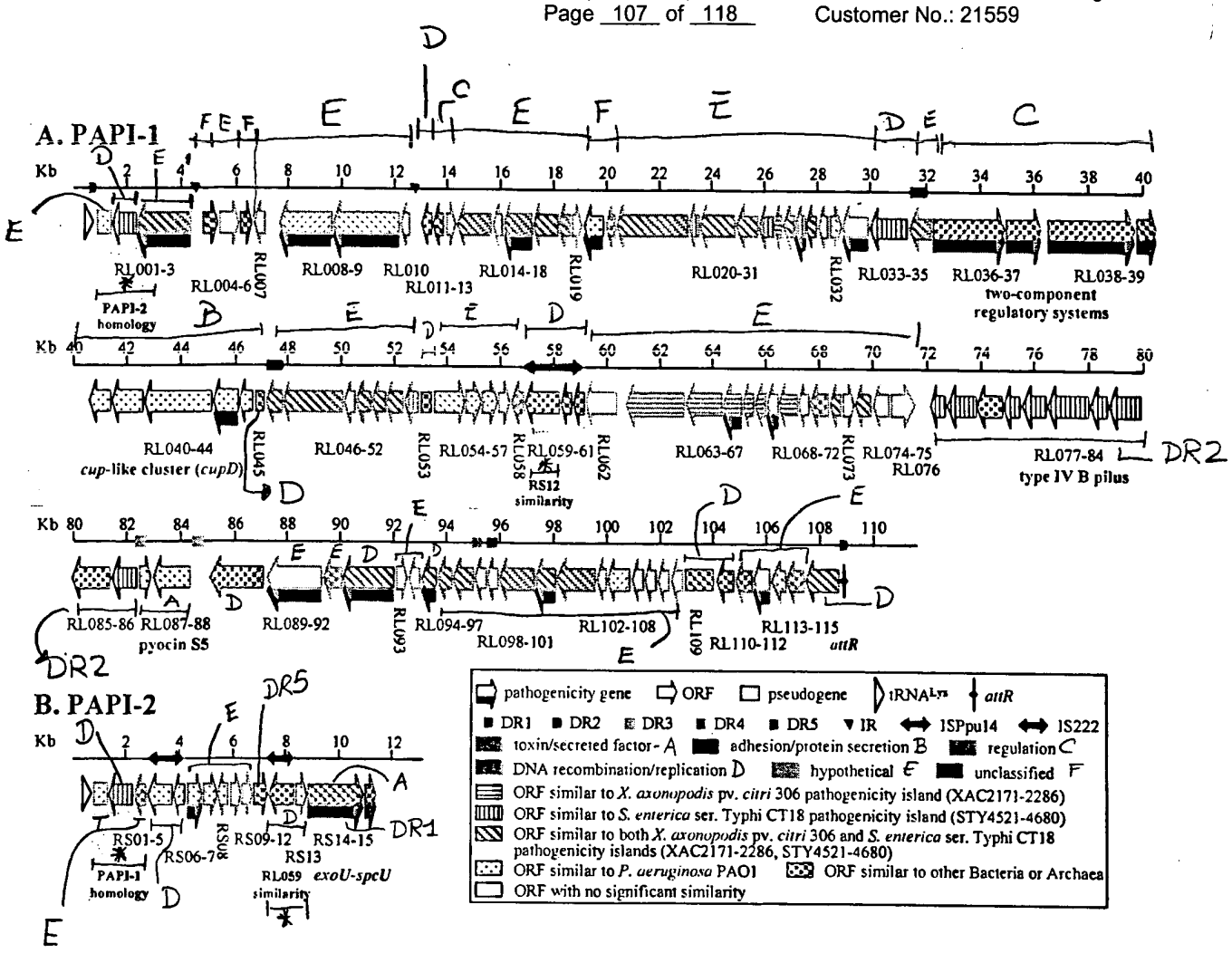
A. PAPI-1
B. PAPI-2
Kb
RL001-3
RL004-6
RL007
RL008-9
RL010
RL011-13
RL014-18
RL019
RL020-31
RL032
RL033-35
RL036-37
RL038-39
two-component regulatory systems
PAPI-2 homology
RL040-44
cup-like cluster (cupD)
RL045
RL046-52
RL053
RL054-57
RL058
RL059-61
RS12 similarity
RL062
RL063-67
RL068-72
RL073
RL074-75
RL076
RL077-84
type IV B pilus
DR2
RL085-86
RL087-88
pyocin S5
RL089-92
RL093
RL094-97
RL098-101
RL102-108
RL109
RL110-112
RL113-115
attR
DR5
DR1
RS01-5
RS06-7
RS08
RS09-12
RS13
RS14-15
RL059 exoU-spcU similarity
PAPI-1 homology
pathogenicity gene
ORF
pseudogene
tRNALys
attR
DR1
DR2
DR3
DR4
DR5
IR
ISPpu14
IS222
toxin/secreted factor - A
adhesion/protein secretion B
regulation C
DNA recombination/replication D
hypothetical E
unclassified F
ORF similar to X. axonopodis pv. citri 306 pathogenicity island (XAC2171-2286)
ORF similar to S. enterica ser. Typhi CT18 pathogenicity island (STY4521-4680)
ORF similar to both X. axonopodis pv. citri 306 and S. enterica ser. Typhi CT18 pathogenicity islands (XAC2171-2286, STY4521-4680)
ORF similar to P. aeruginosa PAO1
ORF similar to other Bacteria or Archaea
ORF with no significant similarity

Figure 60

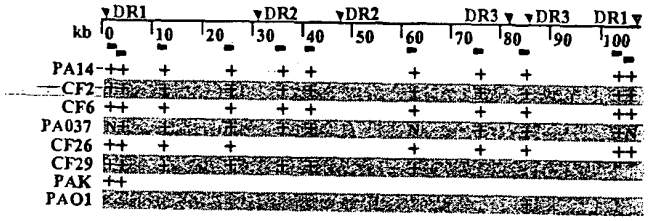
DR1
DR2
DR2
DR3
DR3
DR1
kb
0
10
20
30
40
50
60
70
80
90
100
PA14
CF2
CF6
PA037
CF26
CF29
PAK
PAO1

Figure 61

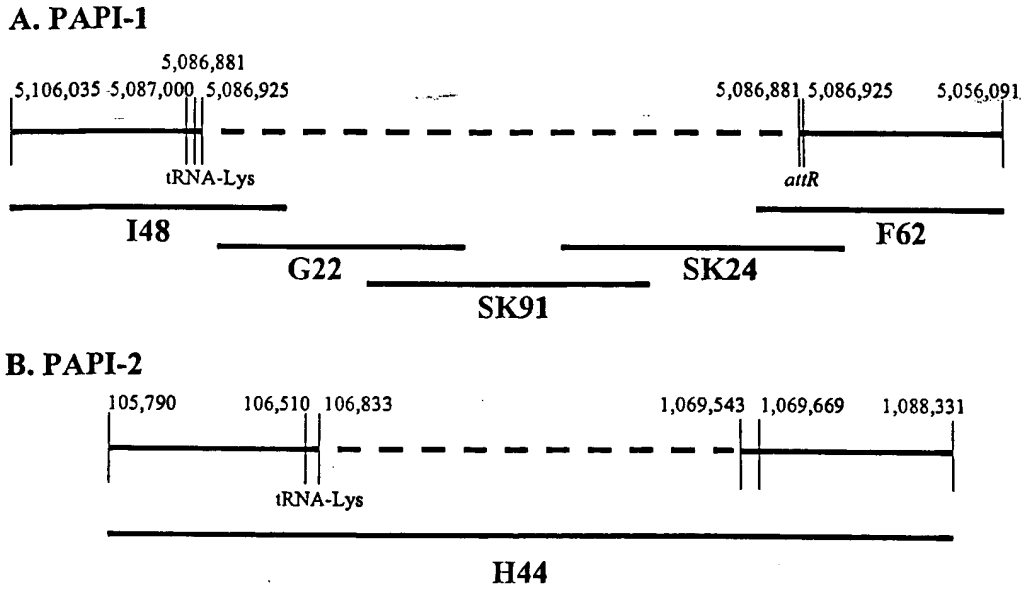
A. PAPI-1
5,086,881
5,106,035
5,087,000
5,086,925
5,086,881
5,086,925
5,056,091
tRNA-Lys
attR
I48
F62
G22
SK24
SK91
B. PAPI-2
105,790
106,510
106,833
1,069,543
1,069,669
1,088,331
tRNA-Lys
H44

Figure 62

FIGURE 63

Strain name*	% Mouse mortality†	Growth in *Arabidopsis* leaf‡	Closest published homologue (organism / GenBank accession no.)
PA14	100	4.9×10^6	
RL003§	**41**	**2.3×10^5**	*P. syringae* pv. *tomato* DC3000 / AAO54371
RL008	**38**	4.1×10^6	*M. acetivorans* C2A / AAM05538 and *P. aeruginosa* PAO1/ AAG05323
RL009	**31**	**1.3×10^4**	*P. aeruginosa* PAO1 / AAG05327
RL016	100	**2.8×10^4**	*P. syringae* pv. *tomato* DC3000 / AAO54383
RL020	**50**	**3.4×10^5**	protein-disulfide isomerase, *P. aeruginosa* PAO1 / AAG04371
RL022	88	3.3×10^6	*P. syringae* pv. *tomato* DC3000 / AAO54394
RL029	**38**	**9.4×10^4**	*P. aeruginosa* C / AAN62148
RL033	**25**	**4.9×10^4**	no significant similarity
RL036	**44**	**1.9×10^5**	two-component sensor *P. aeruginosa* PA14 / AAM15532
RL037	**43**	**1.2×10^5**	two-component regulator *pvrR*, *P. aeruginosa* PA14 AAM15533
RL038	**31**	**4.4×10^4**	two-component sensor *rcsC*, *S. typhimurium* LT2 / AAL21172
RL039	**31**	**2.7×10^5**	two-component regulator *rcsB*, *E. coli* O157:H7 EDL933 / AAG57352
RL043	75	**1.7×10^6**	probable pili assembly chaperone *cupA2*, *P. aeruginosa* PAO1 / AAG05517
RL054	63	NT¶	*P. aeruginosa* PAO1 / AAG05610
RL062	78	NT¶	no significant similarity
RL065	63	**4.5×10^5**	*X. axonopodis* pv. *citri* 306 / AAM37094
RL068	**56**	**2.6×10^5**	no significant similarity
RL090	67	**2.7×10^4**	no significant similarity
RL092	**0**	**1.3×10^5**	topoisomerase I TopA, *X. fastidiosa* 9a5c (plasmid pXF51) / AAF85572
RL095	**50**	5.3×10^5	single-stranded DNA binding protein Ssb, *P. aeruginosa* C / AAN62318
RL101	**38**	1.8×10^6	*Pseudomonas* sp. B13 / CAD60668
RL112	**38**	**1.6×10^4**	no significant similarity
RS06	100	**1.8×10^5**	*P. aeruginosa* PAO1 / AAG04369

	Positions	Length (bp)	Number of identical bp	Genes between the repeats
DR1	744-805 108,700-108,762	63	59	PAPI-1 (108 Kb)
DR2	31,587-32,248 47,100-47,761	662	654	two component regulatory systems and *cup*-like cluster (*cupD*)
DR3	82,574-82,821 85,296-85,540	248	231	pyocin S5 and associated immunity protein
DR4	95,301-95,357 95,358-95,414	57	50	none
DR5	95,767-95,824 95,825-95,881	58	54	none
IR	4,527-4,594 12,825-12,892	68	56	pathogenicity genes and Archaea homologous genes

FIGURE 64

IS name	PAPI-1		PAPI-2		Characteristics of IS	
	Position	Length (bp)	Position	Length (bp)	Original length (bp)	IS family
ISPpu14	56,778-59,119	2,341	7,034-7,999	966	2,383	IS66
IS222	-	-	2,980-4,201	1,222	1,232	IS3

FIGURE 65

FIGURE 66

Function	Prototype name	Type IVB (PAPI-1 in PA14)	Type IVA (PAO1)	Xcp (PAO1)	Hxc (PAO1)	Hpl (PAO1)	Other homologues in PAO1 genome
ATPase	*pulE*	RL082 (*pilQ2*)	*pilB*	*xcpR*	*hxcR*	*hplR*	*hvbA, hxtA*
			pilT				
			pilU				
Peptidase	*pulO*	RL079 (*pilT2*)?	*pilD*	*pilD*	*pilD*	*pilD?*	
Major pilin	*pulG*	RL080 (*pilS2*)	*pilA*	*xcpT*	*hxcT*	*hplT*	
Minor pilin	*pulH*	RL077 (*pilM2*)	*pilE*	*xcpU*	*hxcU*	*hplU*	
	pulI	RL078 (*pilV2*)	*fimU*	*xcpV*	*hxcV*	*hplV*	
	pulJ	RL083 (*pilP2*)	*fimT*	*xcpW*	*hxcW*	*hplW*	
	pulK	RL086 (*pilL2*)		*xcpX*	*hxcX*	*hplX*	
Inner membrane protein	*pulF*	RL081 (*pilR2*)	*pilC*	*xcpS*	*hxcS*	*hplS*	*xqhA*
	PulC	RL084 (*pilO2*)		*xcpP*	*hxcP*		
	PulL			*xcpY*	*hxcY*		
	pulM			*xcpZ*	*hxcZ*		
Secretin	*pulD*	RL085 (*pilN2*)	*pilQ*	*xcpQ*	*hxcQ*		*xqhA, xqhB, xqhC*

FIGURE 67

ORF 7 (SEQ ID NO: 280)

LEFGSATWTRTRDPMINSHLLYRLSYRGTSFFQPWTLPVLLDSRLRGAPFYGCARACQPSDPKSFSSFSTSDKTALPLHAAALSRLPDAHEKAPPKR
GFPCPPPKRSGEDDLVAFHLRRDTGTRREFAGQDQLRQRVLDPALDGPLQRACAIDRVEADGNQLVQRLLAQFQAQLALGQALAQATELDLGDAGDL
LASQRLEHHHFVDPVDEFRTEVRIDRVHHCGTLRLAVAGQLLDLRRTEVGGHHHHGVAEVHRTPVTVGQASVLEHLEENVEYIRMGLLHLVQQHHRV
GLAADRLGQVAAFLEADVARRRADQAGHRVFLHELGHIYPHQRLLGIEEELGQRLAQLGLAHPGRAEEEERAARPVRIGEAGARTAHGVGHGDYRLV
LADHSPMQLLLHAQQLLALALEHLRHRDTGPLGNHFGDFLVGHLVAQQLVLGLAVLVDHLQAAFQVRDGLVLDARHALEVALAPRRLHLLLGLLDLL
LDLRRALHLGLLGLPDLLEVGVFALELDDILLQLGQALPGGFVVFLLQRLALDLQLDQATVETIQFLRLGVDLHADAAGGLVDQVDGLVRQLPIGDV
AVRQLGRGDDRAVGDAHPVVHFIAFLEATEDGDGVFLARFVHQHLLEAALQRGILLDVLAILVEGSSTDAVQLAARQSRLEHVAGVHGTFRLAGADH
GVQFVDEQDDPAFLLAQFVEDRLQAFLELAAELGTGDQRPHVQGQQALVLEAVRHFAVDDALGQALDDGGLADAGFADQHRVVLGPPLQDLDGPADL
VVATDHRVELAFLGALGHVDGVLVQRLARLLDVRVVHRFAATQVGHGILQRLARHALAEQQLAEPGVLVHRGQQYQLAGDELVALLLGQAVSLVEQA
CEILGQVHVAGRALDLRQRVEFFVEAAAQGGDIEADLHQQGLDRTALLLEQGGKQVHRLDGRMVMANGQGLGVGERQLQLAGQTVYSHGSSFLL

ORF7 (SEQ ID NO: 281)

TTGGAATTTGGCTCCGCGACCTGGACTCGAACCAGGGACCCAATGATTAACAGTCATTTGCTCTACCGACTGAGCTATCGCGGAACGTCTTTCTTCC
AACCCTGGACGCTTCCGGTGTTGCTGGATTCGCGTCTCAGAGGCGCGCCATTTTACGGATGCGCGCGGGCATGTCAACCCTCTGATCCAAAAAGTTT
TTCTTCTTTTTCCACGAGCGACAAAACGGCCCTTCCACTGCATGCGGCAGCGCTCTCGCGCCTACCGGACGCCCATGAAAAAGCCCCGCCGAAGCGG
GGCTTTCCCTGTCCGCCCCCGAAGAGGTCAGGCGAAGACGATCTCGTCGCCTTCCACCTTCGCCGAGATACTGGCACCCGGCGCGAATTTGCCGGCC
AGGATCAGTTGCGCCAGCGGGTTCTCGATCCAGCGCTGGATGGCCCGCTTCAGCGGGCGTGCGCCATAGACCGGGTCGAAGCCGACGGCAATCAGCT
TGTCCAGCGCCTCCTGGCTCAGTTCCAGGCTCAGCTCGCGCTCGGCCAGGCGCTTGCGCAGGCGACCGAGCTGGATCTCGGCGATGCCGGCGATCTG
CTCGCGAGCCAGCGGCTCGAACACCACCACTTCGTCGATCCGGTTGATGAATTCCGGACGGAAGTGCGCATTGACCGCGTCCATCACTGCGGCACGT
TGCGCCTCGCGGTCGCCGGCCAGCTCCTGGATCTGCGCCGAACCGAGGTTGGAGGTCATCACCACCACGGTGTTGCGGAAGTCCACCGTACGCCCGT
GACTGTCGGTCAGGCGTCCGTCCTCGAGCACCTGGAGGAGAATGTTGAATACATCCGGATGGGCCTTCTCCACCTCGTCCAGCAGCACCACCGAGTA
GGGCTTGCGGCGGATCGCCTCGGTCAGGTAGCCGCCTTCCTCGAAGCCGACGTAGCCCGGAGGCGCGCCGATCAGGCGGGCCACCGAGTGTTTCTCC
ATGAACTCGGACATATCTATCCGCACCAGCGCCTCCTCGGTATCGAAGAGGAACTCGGCCAGCGCCTTGCACAACTCGGTCTTGCCCACCCCGGTCG
GGCCGAGGAAGAGGAACGAGCCGCTCGGCCGGTTCGGATCGGCGAGGCCGGCGCGCGAACGGCGCACGGCGTTGGACACGGCGACTACCGCCTCGTC
CTGGCCGATCACTCGCCGATGCAGCTCCTGCTCCATGCGCAGCAGCTTCTCGCGCTCGCCCTCGAGCATCTTCGACACCGGGATACCGGTCCACTTG
GAAACCACTTCGGCGATTTCCTCGTCGGTCACCTTGTTGCGCAGCAACTGGTTCTCGGTCTTGCCGTGCTGGTCGACCATCTGCAGGCTGCGTTCCA
GGTCCGGGATGGTCTGGTACTGGATGCGCGCCATGCTCTCGAGGTCGCCCTTGCGCCGCGCCGCCTCCATCTCCTGCTTGGCCTGCTCGATCTTCTG
CTGGATCTGCGCCGAGCCCTGCACCTCGGCCTTCTCGGACTTCCAGATCTCCTCGAGGTCGGCGTATTCGCGCTCGAGCTTGACGATATCCTCCTCC
AGCTTGGCCAGGCGCTTCCTGGTGGCTTCGTCGTCTTCCTTCTTCAGCGCCTCGCGCTCGATCTTCAGCTGGATCAGGCGACGGTCGAGACGATCCA
GTTCCTCCGGCTTGGAGTCGATCTCCATGCGGATGCGGCTGGCGGCCTCGTCGATCAGGTCGATGGCCTTGTCCGGCAGTTGCCGATCGGTGATGTA
GCGGTGCGACAGCTTGGCCGCGGCGATGATCGCGCCGTCGGTGATGCTCACCCCGTGGTGCACTTCATAGCGTTCCTTGAGGCCACGGAGGATGGCG
ATGGTGTCTTCCTCGCTCGGTTCGTCCACCAGCACCTTCTGGAAGCGGCGCTCCAGCGCGGCATCCTTCTCGATGTACTGGCGATACTCGTCGAGGG
TAGTAGCACCGACGCAGTGCAGCTCGCCGCGCGCCAGAGCCGGCTTGAGCATGTTGCCGGCGTCCATGGCACCTTCCGCCTTGCCGGCGCCGACCAT
GGTGTGCAGTTCGTCGATGAACAGGATGACCCGGCCTTCCTGCTTGCCCAGTTCGTTGAGGACCGCCTTCAGGCGTTCCTCGAACTCGCCGCGGAAC
TTGGCACCGGCGATCAGCGCCCCCATGTCCAGGGCCAGCAGGCGCTTGTCCTTGAGGCCGTCCGGCACTTCGCCGTTGATGATGCGCTGGGCCAGGC
CCTCGACGATGGCGGTCTTGCCGACGCCGGGTTCGCCGATCAGCACCGGGTTGTTCTTGGTCCGCCGCTGCAGGACCTGGATGGTCCGGCGGATCTC
GTCGTCGCGACCGATCACCGGGTCGAGCTTGCCTTCCTCGGCGCGCTTGGTCATGTCGACGGTGTACTTGTCCAGCGCCTGGCGCGACTCCTCGACG
TTCGGGTCGTTCACCGCTTCGCCGCCACGCAGGTTGGCCACGGCATTCTCCAGCGCCTTGCGCGACACGCCCTGGCCGAGCAGCAGCTTGCCGAGCC
TGGTGTTCTCGTCCATCGCGGCCAGCAATACCAGCTCGCTGGAGATGAACTGGTCGCCCTTCTGCTGGGCCAGGCGGTCAGCCTGGTTGAGCAGGCG
TGCGAGATCCTGGGACAGGTTCACGTCGCCGGTCGGGCTCTGGATCTTCGGCAGCGCGTCGAGTTCTTTGTTGAGGCCGCTGCGCAGGGCGGCGATA
TCGAAGCCGACCTGCATCAGCAGGGGCTTGATCGAACCGCCTTGCTGCTCGAGCAGGGCGGAAAGCAGGTGCACCGGCTCGATGGCCGGATGGTCAT
GGCCAACGGCCAGGGACTGGGCGTCGGAGAGCGCCAGTTGCAGCTTGCTGGTCAAACGGTCTATTCGCATGGGTCGTCCTTCCTTCTATAG

FIGURE 68 (SEQ ID NO: 121)

ACGTCGGGGGCGCATTGctACGCCTGcAgAATGGTTTCAGGGCCTTAGAAACAGAAAAGCCCACCTaGAC
AGGCGGGCTATTCCATATTGAcATcAcGTCAATGCGGGCCTAATGTTCGGCCCAGACGGCTGCTAGACAA
GAACCGGCGTAACACCCCTTCCTAGCCTATGCAACTCGCCCCGTAGAAAATGGTGGGTCGTGTAGGATTC
GAACCTACGACCAATTGGTTAAAAGCCAACTGCTCTACCGACTGAGCTAACGACCCAAGTATGAGGTGGT
CGGGGTAGAGAGATTCGAACTCCCGACATCCTGCTCCCAAAGCAGGCGCGCTACCGGACTGCGCTATACC
CCGATTGGAATTTGGCTCCGCGACCTGGACTCGAACCAGGGACCCAATG

FIGURE 69

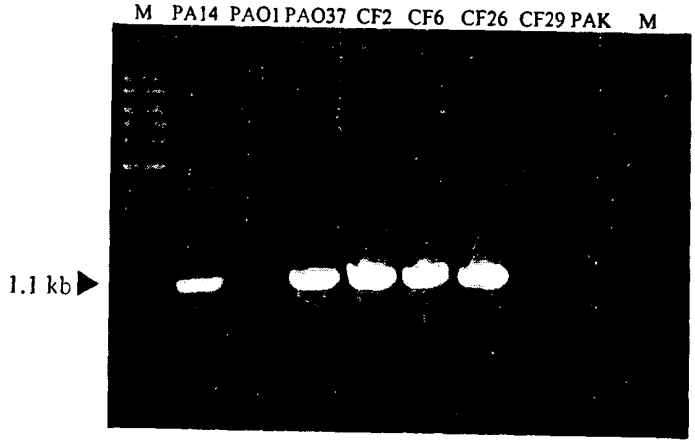

FIGURE 70A
CLUSTAL W (1.82) multiple sequence alignment

```
CF2   (SEQ ID NO: 122)  ---------------------------GATGAAGG-ACCCGAGCGGAACATCCATCTCAC 32
PAO37(SEQ ID NO: 123)  ---------------------------GATGAAGG-ACCCGAGCGGAACATCCATCTCAC 32
CF6   (SEQ ID NO: 124)  ---------------------------GATGAAGGCACCCGAGCGGAACATCCACCTCAC 33
PA14  (SEQ ID NO: 125)  TTTCCGGCACACCCTGGCCACCGACTTGATGAAGGCACCCGAGCGGAACATTCACCTCAC 60
CF26  (SEQ ID NO: 126)  ---------------------------GATGAAGG-ACCCGAGCGGAACATCCACCTCAC 32
                                                   ******** ********************** ** *****

CF2                     GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 92
PAO37                   GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 92
CF6                     GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 93
PA14                    GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACGATGAGCTACATCGAGGCCGACTA 120
CF26                    GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACGATGAGCTACATCGAGGCCGACTA 92
                        ************************************ ************************

CF2                     CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 152
PAO37                   CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 152
CF6                     CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 153
PA14                    CGATCACATGCGTGCCGTGCTGCATGCTAGAAGCCTGGCCCAAGGCGCGCTGGAGAATGT 180
CF26                    CGATCACATGCGTGCCGTGCTGCATGCTAGAAGCCTGGCCCAAGGCGCGCTGGAGAATGT 152
                        *** ********************** ********************* *********** **

CF2                     CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 212
PAO37                   CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 212
CF6                     CAGGAAGGTGGATTACAGCGGCTCCCCGCAGGCCTCTGCCAAACCGAAGCCATGCGGGCA 213
PA14                    CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 240
CF26                    CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAACCATGCGGGCA 212
                        ******************************** ********************* ***********

CF2                     ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCCGGAGGCCAGGACAGAACCTGCAGAACC 272
PAO37                   ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCCGGAGGCCAGGACAGAACCTGCAGAACC 272
CF6                     ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCCGGAGGCCAGGACAGAACCTGCAGAACC 273
PA14                    ACCTCTCGCTCGAGTGAGTGAAGCGCCGCCACCGGAGGCCAGGACAGAGCCTGCAGAACC 300
CF26                    ACCTCTCGCTCGAGTGAGTGAAGCGCCGCCACCGGAAGCCAGGACAGAGCCTGCAGAACC 272
                        ************* ** *******  ***** ***** ************ ************

CF2                     AAGGGAGCACATACCAGGGACAGGCATTCAGGGAGGTCCAACCGTGCGGGAAGAAG---C 329
PAO37                   AAGGGAGCACATACCAGGGACAGGCATTCAGGGAGGTCCAACCGTGCGGGAAGAAG---C 329
CF6                     AAGGGAGCACACACCAGGGACAGGCATTCAGGGAGGTCCAACCGTGCGGGAAGAAG---C 330
PA14                    AAGGGAGCACACGCCAGGGACAGGCATTCAGGGAGGTCCAACCGCGTGGGAAGCAGATGC 360
CF26                    AAGGGAGCACACACCAGGGACAGGCATTCAGGGAGGTCCAACCGAGTGGGAAGCAGAAGC 332
                        ***********  ********************************** * ****** **   *

CF2                     GCTACCACAGCCACCTGACACCTTCGACCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 389
PAO37                   GCTACCACAGCCACCTGACACCTTCGACCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 389
CF6                     GCTACCACAGCCACCTGACACCTTCGATCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 390
PA14                    GCTACCACAGCCACCTGACACCTTCGAACCAAGCGTGCTGTTCACTCTGATGGCTCAAAA 420
CF26                    GCTACCACAGCCACCTGACACCTTCGAGCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 392
                        *************************** * ***************************** *

CF2                     CTTATCGAACCGTGCCGCCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCTGGTGG 449
PAO37                   CTTATCGAACCGTGCCGCCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCTGGTGG 449
CF6                     CTTATCGAACCGTGCCGCCTCGGCATCCGCAGCTCCCGCTGCAACAAGCGGATCTGGTGG 450
PA14                    CTTATCGAACCGTGCCGCCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCAGGCGG 480
CF26                    CTTATCGAACCGTGCCGCCACGACATCTGCGGCTCCCGCCGCAACCAGCGGATCTTGTAG 452
                        ******************* ** **** ** ******** ***** ********  *  *

CF2                     ATGGGGATCTACTGCCCGAAGCAGTCTCGCCTAGCGATACCGATACTGAAGGGCCGGCTA 509
PAO37                   ATGGGGATCTACTGCCCGAAGCAGTCTCGCCTAGCGATACCGATACTGAAGGGCCGGCTA 509
CF6                     ATGGGGATCTACCGCCCGAAGCAGTCTCGCCTAGCGATACCGGTACTGAAGGGCCGGCTA 510
PA14                    ATGGGGATCTGCCGCCCGAAGCAATCTCGCCTAGCGATACCGGTACTGA-GGGCCGGCTA 539
CF26                    ATGGGGATCTGCCGCCCGAAGCAGCCTCGCCTAGCGATACCGGTACTGAGGGGCCGGCTA 512
                        ********** * **********  ****************** ****** **********
```

FIGURE 70B

```
CF2     CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAAGAAAAATCTGGAA 569
PAO37   CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAAGAAAAATCTGGAA 569
CF6     CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAGGAAAAATCTGGAA 570
PA14    CCGGACGAAAGGTAGCCGTGCCTTCCAGCAGATCGTTAGGCCTGTAGGAAAAATCTGGAA 599
CF26    CCAGACGAAAGGTAGCCGCGCCTCCCAGCAGATCGCTGGGCCTGTAGGAAAAATCTGGAA 572
        ** *************** **** ******* *** * ******** ************

CF2     TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAGTTTCACGG 628
PAO37   TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAGTTTCACGG 628
CF6     TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGGGCCC-CGCAATTTCAAGG 629
PA14    TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCAGCGCAATTTCAAGG 659
CF26    TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAATTTCACGG 631
        ****************************************** ***  **** ***** **

CF2     CCAAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCGCATTGCTACGCCTGC 688
PAO37   CCAAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCGCATTGCTACGCCTGC 688
CF6     C-GAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGAGGGCACATTGCTACGCCTGC 688
PA14    CCAATACCACAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCGCATTGCTACGCCTGC 719
CF26    C-AAAACCGCAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCACATTGCTACGCCTGC 690
        *  * *** ****************************** **** ***************

CF2     AGAA-TGGTTTCAGGGCCTTANAAACAGAAAAGCCCACCTTAAATAGGCGGGCTATT-CC 746
PAO37   AGAAATGGTTTCAGGGCCTTAGAAACAGAAAAGCCCACCCTAAATAGGCGGGCTATT-CC 747
CF6     AGAA-TGGTTTCAGAGCCT-GAAAACAGAAAAGNCCACC-TAAATAGGCGGGCTATTTCC 745
PA14    AGAA-TGGTTTCAGGGCCTTAGAAACAGAAAAGCCCACC-TAGAAAGGCGGGCTATT-CC 776
CF26    AGAA-TGGTTTCAGAGCCTTANAAACAGAAAAGCCCACC-TAGATAGGCGGGCTATT-CC 747
        **** ********* ****   *********** ***** ** * ************ **

CF2     ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCAGACGGCTG--CTGG 798
PAO37   ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCANACGGCTG--CTGG 799
CF6     ATATTTGACATCCCG-TCAATGCGGGGCCCTAATGGTTCGGGCCCANACGGCTTGCTTGG 804
PA14    ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCAGACGGCTG--CTAG 828
CF26    ATATT-GACATCACGGTCAATGCGGG--GCTAATGTTC-GGGCCCANACGGNTG---CAA 800
        ***** ****** ** **********   ****** *   ****** **** *
```